

Citation for published version:

Gross, C & Perotti, P 2017, 'Output-Based Measurement of Accounting Comparability: A Survey of Empirical Proxies', *Journal of Accounting Literature*, vol. 39, pp. 1 - 22. <https://doi.org/10.1016/j.acclit.2017.09.002>

DOI:

[10.1016/j.acclit.2017.09.002](https://doi.org/10.1016/j.acclit.2017.09.002)

Publication date:

2017

Document Version

Peer reviewed version

[Link to publication](#)

Publisher Rights

CC BY-NC-ND

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Output-Based Measurement of Accounting Comparability: A Survey of Empirical Proxies

Christian Gross

Institute of Accounting and Control
University of Graz
tel.: +43 316 380 3506
fax: +43 316 380 9565
christian.gross@kelag.at

Pietro Perotti

School of Management
University of Bath
tel.: +44 1225 383 635
fax: +44 1225 386 473
p.perotti@bath.ac.uk

Accepted for publication in the *Journal of Accounting Literature*

Key terms: Comparability, Harmonization, Measurement, Empirical Financial Accounting, Survey

JEL classifications: M40, M41

Acknowledgements: We are thankful for helpful comments from Ralf Ewert, Jamie O'Neill, Thorsten Sellhorn, Alfred Wagenhofer, David Windisch and two anonymous referees. We gratefully acknowledge financial support of the Austrian Science Fund (FWF), P 24911-G11.

Output-Based Measurement of Accounting Comparability: A Survey of Empirical Proxies

Abstract

Accounting comparability has been the subject of significant interest in empirical financial accounting research. Recent literature, particularly that following De Franco et al.'s (2011) influential study, has focused on utilizing the output of the financial reporting process to measure accounting comparability. In this paper, we conduct an early survey of studies using output-based measures of comparability. We provide two distinct contributions to the literature. First, we describe and comment on four important measurement concepts as well as the studies that introduced them. With this methodological contribution, we aim to facilitate the measurement choice for empirical accounting researchers engaged in comparability research. Second, we classify the sub-streams of literature and related studies. In providing this content-related contribution, we sum up what has already been achieved in output-based accounting comparability research and highlight potential areas for prospective research. As a whole, our study attempts to guide empirical researchers who (plan to) undertake studies on accounting comparability in selecting relevant topics and choosing adequate approaches to measurement.

1. Introduction

Accounting comparability¹ is at the forefront of the international standard setters' agenda. Accordingly, it is listed among the desirable properties of financial accounting information in the Conceptual Frameworks of both the International Accounting Standards Board (IASB) and the Financial Accounting Standards Board (FASB). The number of studies on comparability in financial reporting has increased in recent years, particularly after the adoption of the International Financial Reporting Standards (IFRS) in the European Union (EU) and the proposed adoption of the IFRS in the US. Besides these developments in international accounting standard setting, part of this increased interest in accounting comparability also stems from the introduction of new measures in empirical research.

Earlier papers on accounting comparability were based on the comparability of financial reporting inputs (*input-based* approach), i.e., the accounting rules and the choice of reporting methods. Most of these studies derive comparability by counting and weighting differences in accounting method choices over time or across firms. However, recent research has mostly focused on the comparability of the outputs of the financial reporting process (*output-based* approach), most notably of earnings. For example, one of the most widely used output-based measures of comparability is based on the similarity with which accounting data react to economic events. We argue that the output-based approach to measuring accounting comparability entails at least four advantages relative to the input-based approach (see also De Franco et al. 2011): 1) it is more relevant for users because their focus is on the output; 2) it is more objective as it does not require the selection and weighting of the inputs; 3) it is easier to implement in practical terms

¹ What we call 'accounting comparability' in our paper is often called 'financial statement comparability' in the literature. However, we think that the term that we use more accurately describes the underlying concept; in most cases, researchers are not interested in the mere comparability of numbers in the financial statements but in the comparability of the accounting process that leads from economic events to these numbers.

due to the widely available data sources; and 4) it is potentially more accurate in measuring accounting comparability because it allows researchers to control for the similarity of economic events.

The vast majority of contemporary empirical research on accounting comparability employs the output-based approach, which suggests most researchers (implicitly) share our view on the advantages over input-based measurement. In our paper, we provide an early survey of this field of literature. We also extend and complement prior surveys on comparability in, or the harmonization of, financial reporting (e.g., Tay & Parker, 1990; Ali, 2005; Baker & Barbu, 2007a, 2007b), which only examine studies using the input-based approach to measure comparability. Focusing on output-based studies, our contribution to the literature is twofold. First, we describe and critically evaluate contemporaneous ways to measure accounting comparability. Second, we provide an overview of, and propose a classification for, the findings of recent research on accounting comparability.

We provide a methodological contribution by pinpointing differences between output-based comparability measures. First, we focus on the paper by De Franco et al. (2011) in order to describe and discuss their general measurement idea that permeates comparability studies since its publication. While many of the studies that we present directly follow their measurement approach, which is based on the association between earnings and stock prices, the measures in use are often adapted to the respective research settings. In addition to De Franco et al.'s (2011) measurement, we identify, describe, and discuss in detail three other output-based approaches to measuring comparability: Yip and Young (2012) do not only focus on the similarity but also on the dissimilarity of financial reports in order to measure comparability; Bhojraj and Lee (2002) build a measure of comparability by operationalizing ideas from valuation theory; and Kim et al. (2013) consider a comparability measure designed to be relevant for debt instead of equity market par-

ticipants. By providing comparisons and comments on the different measurement ideas and refinements included in these studies, our survey should assist researchers to adequately choose proxies for future studies on accounting comparability.

We also provide a content-related contribution by identifying common themes in studies on output-based accounting comparability and presenting an overview of the research that has recently been conducted in this area. We propose a classification that guides researchers in identifying further research questions regarding the comparability of financial reports. In reviewing the literature, we identify three categories of studies: one stream of research relates comparability to (the introduction of) IFRS; the second group of studies examines the determinants of comparability; and the third line of papers investigates the consequences of comparability. Within each of the three sub-streams of literature, we identify several suggestions for future research.

The remainder of our paper is organized as follows. In section 2, we discuss the meaning of the concept of comparability in accounting, the importance of comparability studies in financial accounting research, the types of comparability measures that exist in the literature, and our focus on output-based measures of accounting comparability; in the last part of the section, we describe and comment on four studies introducing output-based measures of accounting comparability that we consider to be of particular interest to empirical researchers. Section 3 comprises our content-related contribution as we survey and classify recent empirical financial accounting studies that involve output-based measurement, either focusing on comparability as a determinant or a consequence of other concepts. Finally, we summarize our findings and conclude in section 4.

2. The importance of accounting comparability and its measurement in financial accounting research

2.1. The importance of accounting comparability

In this section, we focus on three aspects of accounting comparability. First, we discuss the definition of comparability given by international standard setters. Second, we examine the importance of comparability to different groups of stakeholders. Third, we comment on the historical evolution of the role of accounting comparability over time. All three of these aspects demonstrate the importance of the concept of accounting comparability—each from a different angle.

2.1.1. International standard setters and accounting comparability

Accounting comparability plays an important role in the agenda of both the FASB and the IASB; both standard setters include comparability as a principle and/or qualitative characteristic. To illustrate this, we focus on selected excerpts of the IFRS Conceptual Framework. In this framework, comparability is classified as a qualitative characteristic that enhances—together with other characteristics (verifiability, timeliness, and understandability)—the quality of financial reporting for users of financial statements (IFRS CF.QC4):

‘If financial information is to be useful, it must be relevant and faithfully represent what it purports to represent. The usefulness of financial information is enhanced if it is comparable, verifiable, timely and understandable.’

Later in the IFRS Conceptual Framework, the enhancing qualitative characteristic of comparability is related to or distinguished from two subordinate concepts: ‘consistency’ and ‘uniformity’ (IFRS CF.QC21–23):

‘Comparability is the qualitative characteristic that enables users to identify and understand similarities in, and differences among, items. Unlike the other qualitative characteristics, comparability does not relate to a single item. A comparison requires at least two items.’

‘Consistency, although related to comparability, is not the same. Consistency refers to the use of the same methods for the same items, either from period to period within a reporting entity or in a single period across entities. Comparability is the goal; consistency helps to achieve that goal.’

‘Comparability is not uniformity. For information to be comparable, like things must look alike and different things must look different. Comparability of financial information is not enhanced by making unlike things look alike any more than it is enhanced by making like things look different.’

From the above definitions, some of the dimensions linked to comparability in financial accounting become apparent. First, comparability can be related to single or multiple items in the financial statements; while the former is called ‘consistency’ in the IFRS Conceptual Framework and described as a subordinate concept linked to comparability, the latter is directly referred to as ‘comparability’. Second, comparability can be useful both from a longitudinal perspective on a single firm and a cross-sectional perspective on multiple firms; this can, e.g., also be illustrated by referring to research on comparability in financial accounting: almost all of the empirical studies that we will describe in the further course of this paper use panel data, which means that researchers employ both a longitudinal and a cross-sectional perspective in their research designs. Third, comparability comprises a ‘similarity facet’ as well as a ‘difference facet’, since it both

aims at making like things look alike and different things look differently;² in the IFRS Conceptual Framework this is emphasized by essentially stating that ‘uniformity’—the notion of two financial statement items simply looking alike without further analyzing their nature or composition—is to be differentiated from the concept of comparability.

2.1.2. Accounting comparability and stakeholder needs

Different dimensions of comparability in financial accounting can, e.g., be categorized by looking at different stakeholder needs. First, stakeholders could be interested in the accounting comparability within an organization; such a longitudinal perspective on a single firm often yields the question whether and to which extent items of its financial statements have changed over time. Consistency in a firm’s accounting choices over time, which is understood as a subordinate concept to comparability in the IFRS Conceptual Framework, assures that, e.g., tax authorities can more easily identify discretionary accounting choices made by management to minimize tax payments; auditors, enforcement authorities, and a firm’s board of directors can more easily find mistakes in financial statements by analyzing significant deviations from the current to the last financial report; shareholders are enabled to (re-)allocate their capital by closely monitoring consistently defined profit measures in the income statement over time.

Second, also mentioned in the excerpts from the IFRS Conceptual Framework in section 2.1.1, stakeholders could be interested in comparing a firm to its peers at a given point in time. This cross-sectional perspective is, e.g., important for market intermediaries, such as financial analysts, that compare accounting multiples to make trading recommendations. In addition, auditors and enforcement authorities could also be interested in comparing a firm to its peers, e.g., when rationalizing a firm’s measurement assumptions with respect to estimate-based accounting topics such as impairment of financial or non-financial assets, determining the best estimate for provisions, or model-based (‘level 3’) fair value calculations.

² This last-mentioned distinction between a similarity facet and a difference facet of comparability is used in the research design of Yip & Young (2012) and transferred to a German setting in a recent study by Gross (2016).

Third, stakeholders may be interested in accounting comparability between firms operating in different jurisdictions. This aspect of comparability is important for investors with an internationally diversified portfolio. Despite relevant institutional differences across countries, the same set of accounting standards will be appealing to foreign investors, since familiar accounting rules, principles, and methods are likely perceived as being easier to interpret. Similarly, regulators designing their institutional setting for firms to operate in, are—with markets becoming increasingly international—more likely to introduce accounting standards or, at least, accounting rules comparable to those that have proven to be useful in other jurisdictions or in international standards.

2.1.3. A brief historical perspective on accounting comparability

The relative importance of the multiple dimensions of comparability in financial accounting has evolved over time, with internationalization having an effect on accounting standard setting. Accounting standard setting has evolved from a legalistic approach to an informational approach, its focus transferring from accounting to financial reporting, as Zeff (1993) puts it. This trend has started in the US in 1960 and was observable in the UK, Canada, Australia, and the Netherlands about a decade thereafter and substantially later also in Continental European countries. This development corresponds to international accounting standards—which are typically made with financial markets and the informational needs of shareholders in mind—becoming more important all around the world. Arguably the most prominent international accounting standard setter in recent years has been the IASB, which has evolved as an offspring of professional accounting bodies in nine countries and its predecessor organization, the International Accounting Standards Committee (IASC), to becoming a well-governed standard setting body that basically sets mandatory standards for listed companies in the EU and in many other parts of the world (Zeff 2012). In the meanwhile, financial reports under IFRS are even accepted instead of reports under US GAAP for foreign firms that are cross-listed on US stock exchanges. Hence, the transition to a more informational approach towards accounting standard setting has, together with international standardization, led to higher comparability of financial reporting around the world.

Coming back to the different dimensions of comparability identified earlier, the changing focus of accounting standard setting has drawn attention away from consistency in financial reporting to comparability, particularly across jurisdictions and firms. This corresponds with stakeholders, particularly shareholders, operating in more international environments. The shift from consistency to (international) comparability also becomes apparent when looking at different phases of research on comparability in financial reporting, which has similarly evolved over time and is described in the following section of our paper.

2.2. *Comparability in financial accounting research*

As already mentioned in the last section of this paper, accounting standard setting has moved from a legalistic to an informational emphasis. Similarly, the understanding of accounting comparability in financial accounting research has changed in the last decades. The initial focus on the mere similarity of rules and standards, which is usually referred to as *de jure* (or ‘formal’) harmonization/comparability, has been replaced by a focus on the application of rules and standards, which is labeled *de facto* (or ‘material’) harmonization/comparability. In a period when various local GAAPs were applied across different nations, *de jure* comparability was the somewhat natural focus, since researchers were concerned with a legalistic perspective on the similarity or dissimilarity of rules and standards across different nations. A typical (and well-cited) example of a monograph along these lines is the one edited by Nobes (2001). Since these studies were conducted before international accounting standards have been established as the mandatory reporting basis for consolidated financial reports in many countries, this legalistic perspective has lost most of its relevance nowadays. Rather, the common application of identical standards in countries with different accounting traditions and different institutional environments has become the main subject of interest. Moreover, recent research (e.g., Christensen et al. 2013) emphasizes the role of enforcement activities to understand the economic effect of a (changing) set of accounting standards; *de jure* comparability only focuses on the similarity of the accounting rules and, therefore, this approach overlooks differences in enforcement (or other parts of the institu-

tional environment that firms operate in). Hence, while earlier studies have most often been concerned with *de jure* comparability, virtually all contemporary studies on comparability focus on *de facto* comparability. However, the conceptual differentiation between *de facto* and *de jure* comparability is but one possibility to classify different research foci with respect to accounting comparability. Another vital differentiation, which is based on different measurements of *de facto* comparability, is that between input-based and output-based measures of comparability.

2.2.1. Input-based measurement of de facto comparability in financial accounting research

The conceptual differentiation between *de jure* and *de facto* comparability, which was described in the last section, follows van der Tas (1988) and Tay and Parker (1990). In addition to Tay & Parker (1990), other prior surveys on the comparability of financial reporting are the ones by Ali (2005) and Baker & Barbu (2007a, 2007b). These surveys primarily focus on studies that investigate accounting comparability by looking at accounting method choices, i.e., inputs to the financial reporting process. Comparability is then derived by counting and weighting differences in method choices over time, across firms in one country, or across firms in different countries. Instead of only comparing choices on single accounting items, input-based studies typically weight and aggregate multiple accounting choices to create comparability indices, which are then used in empirical analyses. Recent examples of studies that use such a methodological approach to investigate accounting comparability are, e.g., the studies by Kvaal & Nobes (2010, 2012). They investigate individual accounting choices of firms from different countries to test whether national reporting patterns remain present, even under a common set of accounting standards.

Input-based measurement of accounting comparability comes with the main advantage of addressing individual accounting choices. Policy implications to redraft accounting standards can therefore directly be derived from studies using this form of measurement. However, even though output-based metrics do not come with this very advantage, this measurement approach comes with several other advantages that are discussed in the following section.

2.2.2. Output-based measurement of de facto comparability in financial accounting research

Output-based measurement of *de facto* comparability builds on the outputs of the financial reporting process, mostly earnings. Many studies, following De Franco et al. (2011), measure comparability as the similarity with which accounting data react to economic events. A related approach to measuring accounting comparability is proposed by Yip and Young (2012), who also focus on the dissimilarity of the effect of economic events on accounting numbers. Motivated by valuation theory, Bhojraj and Lee (2002) estimate comparability as the similarity with which price multiples are related to accounting-based fundamental variables. Focusing on the point of view of debt investors, Kim et al. (2013) measure comparability based on the heterogeneity in the adjustments made by Moody's on selected accounting variables. This approach is founded on the premise that, as the heterogeneity of adjustments to reported accounting data decreases, investors are able to make more accurate comparisons among financial statements.

We believe there are at least four advantages associated with using output-based measures of accounting comparability (see also De Franco et al. 2011) relative to an input-based approach. First, financial statement users (as well as most other stakeholders) are typically interested in the outputs of the financial reporting process—such as revenues, the amount of debt, or earnings—and less so in the method choices that lead to these outputs. Second, if input-based indices are used, the selection of the accounting choices and the assignment of the weights in building the indices may be arbitrary. Specifically, a subset of individual accounting choices that are included into the index have to be selected by the researcher; this procedure will very likely include choices that are easy to be individually observed from financial reports, while ignoring choices that are more difficult to track. Moreover, the different accounting choices that are included into the composite index have to be weighted; while equal weights are a common choice, it is not clear whether all accounting are of equal importance indeed. Third, input-based measurement usually limits the sample sizes of the corresponding studies, since researchers have to hand-collect the different method choices from financial reports and cannot use archival data that are readily available in databases. Fourth, input-based measurement of accounting comparability typically ignores the

similarity of economic events between the firms whose financial reports are compared. This means that firms could be classified as being similar to each other even though their seemingly identical accounting method choices were made under different circumstances (see also footnote 3 in section 2.2.2.1 on this argument).

In the further course of this section, we will describe four studies that introduce what we consider to be the main output-based measures of comparability: De Franco et al. (2011; see section 2.2.2.1), Yip and Young (2012; 2.2.2.2), Bhojraj and Lee (2002; 2.2.2.3), and Kim et al. (2013; 2.2.2.4). We focus on these four studies as they each present new measurement approaches, while the other studies that we discuss later in this paper either directly use these measures or modifications of them. In addition to describing the content of each of these papers as well as the comparability measurement used, we comment on the methodological contributions, differences to other measures, and potential concerns.

2.2.2.1. A comparability measure based on the association between accounting outputs and stock returns (De Franco et al. 2011)

When it comes to comparability in empirical financial accounting research based on archival data, the study by De Franco et al. (2011) can arguably be seen as the most influential paper of recent years. They recommend that the way economic events map into earnings be used as an indicator for comparability in accounting practices; this comparability measure is based on the premise that ‘[f]or a given set of economic events, two firms have comparable accounting systems if they produce similar financial statements’ (De Franco et al. 2011, p. 896).³

³ The difference between the two output-based perspectives on comparability—the one solely looking at earnings while the other also considers the economic events that led to the final earnings number—can be illustrated using the following example. Consider two industrial companies with the same earnings number at fiscal year-end. One of these firms experienced a year of notable economic success related to the acquisition of new customers, whereas the other firm lost major clients to its competitors. While the former firm’s shareholders expect an increase in future cash flows and the firm therefore experienced an increase in its share price, the latter firm’s share price declined. While an output-based perspective without control for economic events would suggest that the two reports (or reporting entities) can be seen as being similar (due to the identical earnings number), an approach that controls for the similarity in economic events would identify the mapping of economic events into earnings as being different. In addition to emphasizing

To operationalize the concept of measuring the similarity of reporting practices by looking at earnings and controlling for economic events, De Franco et al. (2011) use stock returns as a proxy for economic events and earnings as a proxy for the financial statement output. Accordingly, they assume that earnings are a linear function of returns and they estimate the parameters of this function through firm-specific time-series regressions. Holding economic events constant for two firms under consideration is then supposed to yield a pairwise comparability score that is not biased by any economic dissimilarity between the two firms under consideration. The pairwise measures between a firm and all its benchmark firms in the same industry are calculated and combined into firm-year-specific summary measures, which are calculated as the mean or median of a firm's comparability with its industry peers.

Having estimated and validated their similarity of accounting functions measure, De Franco et al. (2011) examine the effect of comparability on the number of analysts following the firm and on the properties of the analysts' forecasts. The results show that the probability that a pair of firms is jointly followed by the same analyst is positively associated with the level of comparability between the two firms. The findings also highlight a positive association between analysts' coverage and comparability. These results are consistent with the view that the cost of analyzing a pair of firms decreases in their pairwise comparability. Moreover, comparability is positively associated with analysts' forecast accuracy and negatively associated with forecast dispersion. This evidence is in agreement with the notion that comparability enriches the information environment of a firm.

We highlight the following methodological points and potential limitations of the De Franco et al.'s (2011) output-based similarity of accounting functions measure:

- While the measure captures within-industry comparability, it ignores other aspects of comparability that could also be of interest. If an accounting standards

the importance of controlling for economic events when measuring accounting comparability output-wise, this example also reveals the proxy that is primarily used by De Franco et al. (2011) to do so: changes in share prices.

reform, e.g., abandoned the idea of general purpose financial statements and introduced industry-specific reporting, the comparability scores would rise, even though between-industry comparability would decline. Since De Franco et al. (2011) solely focus on US data, the original measure is also a within-country measure. However, in other studies—for example those by Barth et al. (2012), Yip and Young (2012), and Cascino and Gassen (2015)—cross-country versions are developed.

- De Franco et al. (2011) do not use all of their pairwise comparability scores in their statistical inferences. They use *CompAcct4*, which averages the four highest comparability values for each firm and *CompAcctInd*, which uses the median of all comparability scores for each firm. Hence, while they look at the upper end of the distribution of pairwise comparability scores and at its center, the lower end of the distribution is not examined. As suggested by Yip and Young (2012), distinguishing between the similarity and the difference facet of comparability may be of interest; hence, it could be worthwhile to also examine the distribution's lower end.
- De Franco et al.'s (2011) output-based comparability metric requires the use of data on stock prices. This excludes the possibility of examining unlisted entities and so potentially limits the scope of application of the measure. Moreover, the measure is therefore influenced by return comparability, which could be distinct from accounting comparability, and is affected by differences in stock price efficiency across peer firms, which could be particularly relevant in an international context.⁴
- Importantly, the measure is also affected by economic comparability, i.e., the similarity with which the cash flows of the company react to economic events.

⁴ See, e.g., the related argument by Cascino and Gassen (2015, p. 248), who develop their cash-flow based comparability measures because differing levels of market efficiency could otherwise bias accounting comparability results in their cross-country setting.

Economic comparability is different from accounting comparability as it does not depend on the accounting system. Disentangling accounting and economic comparability can be challenging from a conceptual and empirical perspective, which is why the other output-based measures of accounting comparability can also be similarly criticized.

- DeFranco et al. (2011) implicitly assume that economic comparability is the same for firms belonging to the same industry. However, there may be important differences in economic comparability within an industry in the same accounting period. For example, Srivastava (2016) shows that there are systematic differences in the production functions as well as in accounting and financial characteristics across firms belonging to the same industry at a given time; these differences are due to the fact that a new cohort of firms entering an industry uses higher amounts of intangible inputs than incumbent firms.⁵ The other output-based measures are also based, in some aspects, on comparisons performed within industry-years and are, therefore, subject to similar criticism.

2.2.2.2. *Comparability measures based on the similarity and dissimilarity of financial reports* (Yip & Young 2012)

Yip and Young's (2012) study builds on the output-based accounting comparability measurement introduced by De Franco et al. (2011); however, they refine the measurement of the construct under consideration by emphasizing that the increased similarity of similar firms as well as the *decreased* similarity of dissimilar firms can both increase overall accounting comparability in the cross-section. Relying on quotes from the FASB and the IASB, they separate the similarity facet inherent in comparability from a difference facet and state that comparable accounting standards intend to make '[...] similar things look more alike *without making different things look less different*' (Yip & Young 2012, abstract, emphasis added). Moreover, they separate within-country

⁵ On a related note, the accuracy of industry classification is a controversial issue (e.g., Hoberg & Philips 2016).

comparability from between-country comparability. Yip and Young (2012) use three different measures of accounting comparability—including a modified version of De Franco et al.'s (2011) measure—on listed firms from 17 European countries that mandated IFRS reporting in consolidated financial reports from 2005 onwards. To assess both facets of accounting comparability, they use each of the measures on variations of different and similar cross- and within-country firms, with similarity in the matching being based on industry affiliation. Interestingly, the results of the study show that similar firms became more similar across countries after IFRS adoption, whereas no consistent results on the difference facet or within-country comparability are found.

Other than a modified version of De Franco et al.'s (2011) measure, Yip and Young (2012) use a 'degree of information transfer' as well as 'similarity of the information content of earnings (ICE) and book value of equity (ICBV)' measures. The first step toward estimating the degree of information transfer is to compute the abnormal stock returns for firms releasing earnings announcements. Second, the mean cumulative abnormal return of all matched non-announcing firms is calculated for the announcement period of the respective announcing firm. The matching of announcing firms to non-announcing firms is conducted separately for the similarity and the difference facet tests and for the cross- and within-country analyses. Third, the average absolute value of the cumulative abnormal returns on the non-announcing firms is regressed on the absolute value of the cumulative abnormal return per share of the announcing firm, some control variables and some interaction terms. The independent variable of interest in the separate within- and cross-country regressions for similar and different firms is the interaction term of an IFRS dummy variable, indicating the affiliation of any earnings announcement to the pre- or post-IFRS period, and the cumulative abnormal return of the announcing firm. In pooled regression analyses, further distinctions between the different comparability facets as well as within- and between-country firms are made by using respective three-way interaction terms as variables of interest.

The similarity of the ICE and ICBV employed by Yip and Young (2012) as a third comparability measure is based on the Ohlson (1995) model. It is either estimated for different industries within a country or for different countries while holding industries constant. In this model, the effect of net income, the book value of equity, a set of dummy variables indicating country or industry affiliation, the interaction terms between the industry/country dummy and net income, and the interaction between the dummy and the book value of equity on the market value of equity is estimated in a linear regression model. The coefficients of interest in these regressions are the ones on the interaction between net income and the industry/country dummy (ICE) and between the book value of equity and the dummy variable (ICBV). To examine the similarity facet between countries, Yip and Young (2012) estimate the model within all the different industries (with enough available data) and every possible two-country combination in the sample. Insignificant (significant) ICEs and ICBVs are then assigned to a comparability score of 1 (0). Moreover, to test the difference facet within countries, they estimate the Ohlson (1995) model using a set of firms from the service industry and another set from the manufacturing industry within each country. Due to the restriction imposed by the limit of one dummy variable in the model, Yip and Young (2012) can neither examine the difference facet across countries nor the similarity facet within countries for this third comparability measure.

We believe that the following methodological points and potential limitations should be considered in the interpretation of Yip & Young's (2012) paper:

- While De Franco et al. (2011) employ a long timeline of US data, Yip and Young (2012) examine comparability around the introduction of the EU's IFRS mandate. Hence, while De Franco et al. (2011) operate without an exogenous event that separates their dataset, Yip and Young have a pre- and a post-IFRS period. In this setting, a further restriction of De Franco et al.'s (2011) measure becomes obvious. In studies including an event that splits the sample period, the necessity to calculate clean measures in the pre- and the post-period in a timely fashion is

at odds with the procedure for calculating a comparability score that is based on data from the previous 16 quarters (four years). Since Yip and Young's (2012) post-period is limited to three years of data and due to the use of semiannual date, their sample remains relatively small.

- While De Franco et al. (2011) calculate *de facto* comparability for all available firm pairs within one industry, Yip and Young (2012) only compare previously matched firm pairs. The similarity facet is examined by using comparability metrics on a matched sample of firms from identical industries that are similar in terms of the magnitude of their total assets. For testing the difference facet, Yip and Young's (2012) comparability scores are computed for matched firms with a similar magnitude of total assets but operating in dissimilar industries. Although the comparability score calculation in similar and dissimilar industries is an original idea that enables interesting sample splits, it is noteworthy that this research design choice entails the implicit assumption that not only the comparability scores but also the industry classifications in use measure similarity in financial reporting.⁶ Since some industry classifications seem to yield better results in relative firm valuations than others (Bhojraj et al. 2003), this assumption can be questioned.
- Yip & Young (2012) also separate a within-country from a cross-country dimension. For the cross-country analyses, firms from similar or dissimilar industries are matched on the basis of total assets with both firms coming from different countries. Even though this is an interesting aspect of Yip and Young's (2012) study, their cross-country perspective and the general selection of countries may

⁶ This assumption is also implicit in De Franco et al.'s (2011) comparability score calculation. However, De Franco et al. (2011) use the industry classification for pre-selecting *a set of* industry peers; for every firm pair belonging to the same industry group, De Franco et al. (2011) calculate a comparability score in the first step; in the second step, aggregate comparability metrics are derived. In contrast, Yip & Young (2012) use a one-to-one matching (that is based on total assets similarity and industry affiliation) that leads to *one* industry peer for each firm; these firm pairs are then used for the comparability score calculation. Thus, the industry classification in Yip & Young's (2012) study is arguably more central to their findings.

be contentious. As Cascino and Gassen (2015, footnote 25, p. 272) point out, no firms from benchmark countries, i.e., countries that did *not* adopt IFRS, are included in the sample. By limiting their sample to IFRS-adopting countries, Yip and Young's (2012) results could be attributable to general time trends that are present independent of IFRS adoption. To account for this research design issue, Cascino and Gassen (2015) conduct similar analyses by examining firms from 29 countries, only 14 of which have required IFRS reporting within the sample period. Firms from the other 15 countries are employed as a control group. Interestingly, Cascino and Gassen's (2015) results call those found by Yip and Young (2012) into question, since they only find weak comparability effects in their first analyses that are—apart from the sample composition—very similar to the analyses by Yip and Young (2012).

- Besides the modified measure from De Franco et al. (2011), Yip and Young (2012) use a measure on the degree of information transfer as a proxy for comparability. It measures the effect that surprises in earnings announcements of announcing firms have on the stock returns of non-announcing firms. It is examined through linear regression models—again using four samples: two within-country samples with firms from similar or different industries and two between-country samples for the similarity and the difference facet. Unlike other comparability proxies, this computation procedure yields coefficient estimates for each of the four industry-country classifications and not a firm- or industry-specific metric. Due to this different aggregation level that limits the applications for these comparability proxies, the degree of information transfer analysis is more suited to complement than to replace other comparability analyses on the firm- or industry-level. Moreover, an important challenge in the interpretation of this measure is understanding what portion of it is affected by economic comparability across firms in the same industry.

- Yip and Young (2012) also use the similarity of the ICE and ICBV based on Ohlson (1995) to examine comparability in their setting. In this model, firms' market values are regressed on net income, the book value of equity, an industry or a country indicator, and the interaction of the respective indicator with net income and the equity book value. If the coefficient on the first interaction term (with net income) is insignificant (significant), Yip and Young (2012) assign an ICE comparability core of 1 (0). If the coefficient on the second interaction term (with equity book value) is insignificant (significant), the value on the ICBV comparability score is assigned as 1 (0). This focus on the insignificance of the two coefficients is theoretically comprehensible: a significant coefficient would indicate that firms from different sets of countries/industries have a different ICE/ICBV; however, it may be noted that focusing on the insignificance of regression coefficients in a small sample setting is unlikely to yield robust indicators of comparability.
- Similar to De Franco et al.'s (2011) measurement approach, Yip and Young's (2012) measures are based on market data; therefore, they are subject to the same potential concerns related to market data that are already discussed in section 2.2.2.1.

2.2.2.3. A comparability measure based on valuation theory (Bhojraj & Lee 2002)

Bhojraj and Lee (2002) present a method for the selection of comparable firms based on valuation theory and applied to accounting multiples. The method is designed to improve analysts' and researchers' selection of comparable firms. Bhojraj and Lee (2002) refer to their approach to identifying comparable firms as the 'warranted multiple method'. Two widely used reference multiples are considered: the price-to-book ratio and the enterprise-value-to-sales ratio. The warranted multiples are obtained as the fitted values of yearly cross-sectional regressions using nine explanatory variables (which measure profitability, growth, and risk) on these reference multiples. In Bhojraj and Lee's (2002) application with US data, the cross-sectional regressions employ all

firms with relevant data available at the intersection of Compustat, the Center for Research in Security Prices (CRSP), and the Institutional Brokers' Estimate System (I/B/E/S).

The method of warranted multiples is motivated by valuation theory. Specifically, the residual income model can be used to obtain an expression of the price-to-book ratio as a function of the cost of equity capital, current book value of equity, and expectations on the future return on equity and on the future book value. Similarly, the residual income model provides an expression for the enterprise-value-to-sales ratio as a function of the cost of capital, current total sales, and expectations on future operating profit margin and on the payout ratio. Bhojraj and Lee (2002) argue that the explanatory variables chosen for the cross-sectional regressions approximate the determinants of the price-to-book and the enterprise-value-to-sales ratio identified by the residual income model.

Bhojraj and Lee (2002) begin their empirical proceedings by estimating the respective cross-sectional regressions, using the explanatory variables based on profitability, growth, and risk to explain the price-to-book or the enterprise-value-to-sales ratio, respectively. For each firm, last year's coefficient estimates are then used together with the firm's current profitability, growth, and risk variables to predict the warranted multiples for each firm-year in the sample (separately for the price-to-book and the enterprise-value-to-sales ratio). Having predicted warranted multiples for each sample firm and year allows matching firms based on the similarity in these multiples: the comparable firms are those firms with warranted multiples closest to that of the target firm.

The results indicate that the warranted multiples method strongly outperforms standard matching methods that are often based only on similarity in size and industry. An out-of-sample validation of the method compares the explanatory power of models relating future price-to-book and enterprise-value-to-sales ratios to a set of *ex ante* measures based on alternative definitions

of comparable firms. The incremental benefit of the warranted multiple approach is more pronounced for stocks belonging to ‘new economy’ industries (i.e., firms from the tech, biotech, and telecommunication sectors).

The following methodological points and potential limitations should be considered about the warranted multiples method:

- Bhojraj and Lee’s (2002) analysis is mainly aimed at improving valuation techniques and providing a methodological contribution for control-sample choices of empirical researchers. However, the methods can be used to measure the degree of pairwise comparability and, more generally, the overall comparability of firms’ accounting systems. To obtain a firm-year specific measure, one could, e.g., calculate the pairwise absolute differences in a warranted multiple between a firm and its industry peers and then compute a firm-specific comparability summary measure.
- To validate their method, Bhojraj and Lee (2002) investigate the predictive ability of the warranted multiples with respect to actual future multiples. The results show a substantial improvement in the predictive ability relative to traditional matching methods that are based on only industry and size. These results strongly support the relevance of the warranted multiple method. A potentially interesting issue that could be addressed by future research would be the examination of the predictive ability of the more widely used of De Franco et al.’s (2011) measure with respect to future multiples.
- The warranted multiples method is specifically designed for equity investors to use as a stock picking tool. As a comparability measure, the method is therefore also primarily relevant for equity investors, since it relates the stock price variation to the variation in reported accounting numbers—both when focusing on the price-to-book ratio and the enterprise-value-sales ratio. However, the enterprise

value is the sum of the debt and equity values, which makes the measure—at least when focusing on the enterprise-value-sales ratio—also relevant for debt holders. Hence, interpreted in a comparability setting, Bhojraj and Lee’s (2002) measurement combines the foci of both De Franco et al. (2011) and Kim et al. (2013) in that it neither exclusively focuses on equity nor on debt market participants.

- The data necessary to estimate the explanatory variables is widely available in both the US market and the more developed international markets. While the application presented by Bhojraj and Lee (2002) concentrates on the US market, the method can also be used to assess cross-country comparability. For example, Young and Zeng (2015) employ the warranted multiples method in a cross-country setting. They use the warranted multiple framework to show how higher comparability leads to an improved selection of international peer firms and greater valuation accuracy.
- A possible limitation of the measure is that it requires analyst data for one of the explanatory variables (consensus forecast on long-term growth). This excludes the smallest firms from the analysis. Importantly, prior research (e.g., De Franco et al. 2011) shows that analyst forecast behavior is associated with comparability. Hence, the focus on firms with analyst coverage may lead to sample selection problems for comparability studies which employ this measure.
- To obtain the multiples, market data is used. Thus, the potential concerns with respect to market data that were already discussed in section 2.2.2.1 also apply to the warranted multiples method.

2.2.2.4. A comparability measure for debt market participants (Kim et al. 2013)

In contrast to the previously described studies, Kim et al. (2013) propose two measures of comparability specifically designed to be relevant for debt market participants and to examine the role of comparability in debt markets. The measures are based on a database compiled by Moody’s—‘Financial Metrics’—which provides adjusted financial accounting data for the purpose of rating

valuation. Moody's adjusts the financial statements in Financial Metrics in order to '[...] improve the comparability of financial statements' (Kim et al. 2013, p. 788). The measures are defined as the negative value of the variability of Moody's adjustments affecting the interest-coverage ratio or of the adjustments for non-recurring income items within an industry-peer group. These measures are meant to capture the heterogeneity of the adjustments; a lower heterogeneity is assumed to indicate higher comparability.

The analysis focuses on the adjustments to the interest-coverage ratio as well as the adjustments for non-recurring income items because they are argued to summarize all of Moody's adjustments regarding solvency and profitability. Furthermore, the two sets of adjustments are expected to be uncorrelated, which implies that the two measures complement each other and so can be jointly used in a multivariate analysis on the consequences of comparability without causing multicollinearity issues. The rationale for interpreting these measures as proxies for comparability is that if the heterogeneity of adjustments to reported accounting data decreases, bond investors are able to make less judgmental comparisons among financial statements within an industry. Therefore, comparability is interpreted to decrease with the heterogeneity of these adjustments.

The results of the analyses by Kim et al. (2013) show that their measures of comparability are positively associated with bond liquidity, as measured by the bid-ask spread of traded bonds. This lends support to the view that comparability helps in reducing information asymmetries—also in debt markets. In particular, Kim et al. (2013, p. 785) argue that comparability enables '[...] less informed investors to conduct simple and standardized but still effective financial analyses'. At the same time, comparability is negatively associated with the credit spreads of bonds and credit default swaps. These findings are interpreted as providing evidence for information asymmetry being reflected in prices. As bond and credit default swap markets are less liquid and so less competitive than equity markets, Kim et al. (2013) note that the characteristics of these debt markets under consideration should increase the strength of the association between prices and

information asymmetry relative to equity markets. Finally, comparability is positively associated with the steepness of the credit default swap term structure between one and five years. Since the steepness of the term structure may be interpreted as being negatively related to uncertainty regarding the default probability of a firm, the results suggest a negative association between comparability and uncertainty for debt market participants.

We identify the following methodological points and potential limitations in the paper by Kim et al. (2013):

- Kim et al.'s (2013) main contribution lies on their focus on the debt market. Moody's is a well-informed and important bond-market participant, which ensures that the focus on its valuation adjustments is relevant for bond investors. A related question is whether the measures could also be relevant for equity investors. Since Moody's adjustments involve a group of financial statement line items that are related to profitability, Kim et al. (2013) argue that the measures may also be of interest to equity investors. Future research could investigate whether this presumption holds true.
- A similar approach to that of Kim et al. (2013) could be used to develop comparability measures in any situation where equity analysts report accounting numbers which are based on GAAP numbers but need to be adjusted. For example, equity analysts often exclude some (mostly non-recurring) items when forecasting earnings (with the result being sometimes referred to as 'street earnings'), which is reflected by the data reported in analyst-tracking services (e.g., I/B/E/S). We believe that examining these adjustments may be an interesting venue for future research to develop new comparability measures mainly relevant to equity holders.⁷

⁷ We thank an anonymous reviewer for this suggestion.

- A possible limitation of Kim et al.'s (2013) metric is that it is only available for firms with publicly traded bonds. This constrains their sample to larger firms. If size was systematically related to the determinants or consequences of comparability—for example if smaller firms were those for which the comparability benefits would be greatest—such a sample limitation could pose a problem in empirical research examining comparability as an independent or dependent variable.
- Unlike De Franco et al.'s (2011), Yip and Young's (2012), and Bhojraj and Lee's measures, Kim et al.'s (2013) measurement approach can also be used for firms with unlisted equity. This could be advantageous as such firms represent an area only scantily explored by comparability research thus far. However, since their measures require firms to instead have publicly traded bonds, and only few firms have publicly traded bonds without simultaneously having listed equity, this advantage remains a rather theoretical one.
- A further potential problem of Kim et al.'s (2013) measure is linked to its aggregation level. The variability of the adjustments is estimated within a quarter-industry group; therefore, the measures are calculated for industry quarters rather than firm quarters. Hence, Kim et al.'s (2013) methodology cannot be directly used to estimate pairwise comparability. We suggest two possible ways through which the methodology could be extended to obtain firm-specific metrics. First, researchers could consider the magnitude of the adjustments at the firm level instead of their variability at an industry level. Even if this approach is practicable, it would not control for different industry-adjustment levels any longer, which would make it more challenging to disentangle accounting comparability from economic similarity when calculating comparability scores. However, this could be countered by using matched samples of similar companies, for example by applying propensity score matching. Second, one could alternatively consider

the variability at the firm level over a period of time, for example over the last four quarters, to obtain a firm-year specific metric. Again, this approach would not allow researchers to control for industry effects. However, researchers could try to observe changes in credit ratings for matched firm pairs in a longitudinal setting.

- One key feature of Kim et al.'s (2013) measure is that it does not use equity market data. On the one hand, this is a disadvantage, since it implies that the measure exploits a narrower information set than the market-based measures. On the other hand, not using market data can be an advantage, as the potential problems linked to return volatility and market efficiency that we discussed in section 2.2.2.1 do not apply.

3. Main findings of the literature on output-based accounting comparability and future research directions

While section 2 of this paper is concerned with studies that are influential in terms of their measurement of the comparability between reporting entities, this section identifies different streams of literature related to the output-based comparability of financial reports. Since many articles that we discuss in this section are related to comparability effects in the wake of IFRS adoption, particularly in the EU, we deal with these studies in section 3.1. Other studies related to comparability but not primarily related to IFRS are then dealt with in section 3.2. Within these latter studies, we identify papers that examine the determinants and others that consider the consequences of accounting comparability. Within each sub-stream of literature, we provide several suggestions for future research.

3.1. IFRS and comparability

After the mandatory IFRS reporting regime for listed firms in the EU became effective in 2005, the between-entity aspect of comparability across different countries was expected to change,

since firms previously reporting under different local GAAPs were now confronted with a common set of accounting standards for their consolidated reports. This expectation is also inherent in the preamble (para. 5) to the European Commission's (2002) International Accounting Standards (IAS) regulation, where the convergence of accounting standards in Europe is deemed '[...] important for the competitiveness of Community capital markets [...]' and advantages are expected '[...] for cross-border transactions or listing anywhere in the world.' Linking *de jure* to *de facto* comparability in this context is particularly interesting, since some literature exists that observes divergent reporting patterns even after firms have started to report under common standards (Daske et al. 2013; Kvaal and Nobes 2010 & 2012). From a comparability point of view, these studies indicate that increased *de jure* harmonization does not automatically lead to increased *de facto* comparability. Besides the literature on the cross-country comparability of entities in countries where IFRS reporting requirements have been introduced, the comparability of IFRS and US GAAP is a recurring theme in studies on *de facto* comparability and IFRS reporting. In this regard, it remains a particularly interesting question whether the US should allow IFRS as a voluntary reporting alternative to US GAAP in the US capital market—not only for cross-listed firms but also for US firms.

In structuring this section of our survey, we differentiate between studies that try to identify an IFRS adoption effect on comparability (see section 3.1.1) and several other IFRS-related issues linked to comparability (see section 3.1.2). All of the studies on the association between accounting comparability and IFRS that we review are listed and briefly described in Table 1.

< Insert Table 1 about here >

3.1.1. The effect that IFRS adoption has on comparability

Focusing on the *de facto* comparability of IFRS and US GAAP firms after firms from IFRS-adopting countries adopted IFRS, Barth et al. (2012) investigate the extent to which international IFRS adopters report similarly to US GAAP firms. The analysis is based on two measures of

comparability which are obtained as modifications of the measure of DeFranco et al. (2011).⁸ The first measure is aimed at capturing the similarity of accounting systems whereas the second measure is concerned with the similarity in the value relevance of accounting information. Using both measures, Barth et al. (2012) find evidence consistent with the notion that comparability between IFRS and US GAAP firms increased after IFRS firms adopted IFRS. The effect is stronger when IFRS adoption is mandatory, in common law countries and where enforcement is stricter.

Using both of the accounting comparability measures developed by Barth et al. (2012) and an international sample of firms domiciled in 27 different countries, Barth et al. (2013) test whether the voluntary IFRS adoption of former local GAAP firms enhances the comparability with firms that had already adopted IFRS and reduces the comparability with non-adopting firms in the same country. They conduct this study to provide indirect evidence relevant to the US Security and Exchange Commission's decision to potentially allow for voluntary IFRS adoption by firms domiciled in the US instead of reporting under US GAAP. While the comparability measures in use equal those employed in Barth et al. (2012), the voluntary adoption setting of Barth et al. (2013) allows the construction of two different matched samples, one comprising 'adopters' and 'adopted' firms and the other one containing 'adopters' and 'non-adopters'.⁹ Using

⁸ The accounting system comparability measure is calculated in three steps. In contrast to the approach of De Franco et al. (2011), Barth et al. (2012) do not estimate longitudinal firm-level regressions in their first step. Rather, they estimate separate panel regressions for US GAAP and IFRS firms in the pre-(IFRS-) and the post-adoption period to calculate within-sample fitted prices, returns, and cash flows. In the second step, the regression coefficient on the accounting amount proxy of one set of firms is used to predict the between-sample fitted prices, returns, and cash flows of the other set of firms and vice versa. In the third step, the absolute differences of within and between sample fitted cash flows are averaged over matched firm pairs and then defined to approximate the accounting system comparability. Value relevance comparability, i.e., a second proxy that Barth et al. (2012) use to measure comparability, is also calculated in three steps. The first step is to estimate nested regressions solely containing country- and year-fixed effects on the right-hand side and the respective economic outcome variable on the left-hand side. In the second step, the effects of combinations of net income, equity book value, and loss dummies as well as country and year fixed effects on the outcome variables are estimated. In the final step, the adjusted R^2 differences between the nested and the full regressions are calculated for the pre- and post-period for IFRS and for US GAAP firms, respectively. Then, the R^2 differences between the US GAAP and IFRS firms are computed.

⁹ 'Adopters' are defined as firms that did not report under IFRS in the pre-adoption period and initially started doing so in the post-adoption period. Thus, the labelling of the pre- and post-adoption periods follows the accounting standards regime used by these adopters. 'Adopted' firms are, in turn, defined as firms that had already voluntarily adopted IFRS in the pre-adoption period of adopter firms and kept reporting under IFRS in the adopters' post-adoption period. Finally, 'non-adopters' are defined as firms that had local GAAP adopted in both the pre- and the post-adoption period of adopters.

both similarity and dissimilarity hypotheses and different matched samples to test them, Barth et al. (2013) implicitly build on ideas from Yip and Young (2012), who nevertheless operationalize these ideas differently. As hypothesized, Barth et al. (2013) find that IFRS adoption is associated with the accounting amounts of adopters becoming more similar to those of adopted firms and less similar to those of non-adopting firms. Aside from their comparability results, they also find that adopters generally exhibit increases in liquidity, share turnover, and stock price synchronicity after IFRS adoption.

Caban-Garcia and He (2013) provide a contribution on the effect of IFRS adoption on *de facto* accounting comparability, focusing on Denmark, Finland, Sweden, and Norway. In 2005, all of these countries required listed firms to report under IFRS. Additionally, the Copenhagen, Helsinki, and Stockholm stock exchanges were consolidated into the OMX Nordic Exchange in the very same year. However, Norway was not involved in this stock exchange merger, which allows Caban-Garcia and He (2013, p. 72) '[...] to separate the effect of harmonized accounting standards from that of harmonized regulation.' They find that comparability for countries in the sample generally improved from the 2001–2004 to the 2005–2008 period (even if they document mixed results for Finland) and that the comparability improvements are especially pronounced for firms in the countries that were involved in the stock exchange merger (in comparison to the benchmark country Norway). For their comparability measurement, earnings-price ratios (similar to Land & Lang, 2002) as well as the similarity of accounting functions measure by De Franco et al. (2011) are used.

In contrast with the studies discussed above, Lang et al. (2010) find that accounting comparability decreases after IFRS adoption. They examine a set of firms from 23 countries that adopted IFRS together with firms from 23 non-adopting countries. The analysis is motivated by the premise that comparability may not be desirable if it implies that dissimilar events are treated similarly. Lang et al. (2010) concentrate on cross-country earnings comovement and cross-country accounting comparability, with the latter concept being estimated by modifying De Franco et

al.'s (2011) similarity of accounting functions measures to be applicable in an international setting. They measure the quality of the information environment by examining analysts' coverage, analysts' forecast accuracy, analysts' forecast dispersion, and bid-ask spreads. The results presented by Lang et al. (2010) show that cross-country earnings comovement is negatively associated with the quality of the information environment. Interestingly, this finding is contrary to the results on earnings comovement in a single country that De Franco et al. (2011) present for the US. However, consistent with findings from prior literature, cross-country accounting comparability is positively associated with the quality of the information environment. Furthermore, difference-in-differences tests indicate that IFRS adoption led to an increase in cross-country earnings comovement, although no significant change in accounting comparability occurred. Accordingly, Lang et al. (2010) claim that the increase in earnings comovement led to a decrease in the quality of the information environment.

Cascino and Gassen (2015) complement the literature on the comparability of accounting standards by emphasizing the moderating effect that compliance has on the relation between IFRS adaption and accounting comparability. First, using two modified versions of De Franco et al.'s (2011) similarity of accounting functions measure that are suitable for their international setting with firms from 29 different countries, they find that comparability increases after IFRS adoption are marginal in magnitude and do not persist across different model specifications. Second, in the search for a reason that explains the lack of more convincing evidence in their first analysis, they examine a hand-collected sample of financial reports from Italian and German firms that were filed in 2006, since these two countries are similar in some aspects (e.g., macroeconomic size and distance from local GAAP to IFRS), but different in other aspects that are related to financial reporting compliance (e.g., corporate governance and investor protection). They compute a compliance score based on the evaluation of measurement and disclosure compliance of IFRS 2 as well as IAS 11, 17, 19, 36, 38, and 39. In examining its determinants, some significant cross-country differences are observable. In their third analysis, they measure compliance as the first component of a principal component analysis of three corporate governance variables (auditor

quality, board independence, and government ownership) that are available for the large international sample that was also used in their first tests. Using a multivariate regression analysis, which examines the moderating effect of compliance on the association between IFRS adoption and comparability, IFRS adoption is found to bring about increases in comparability for firms with distinct compliance incentives.

Liao et al. (2012) examine the cross-country comparability of German and French firms in the first three years following IFRS adoption. They measure comparability as the ‘usefulness’ of earnings and book values in explaining the variation in stock prices. Specifically, following a model proposed by Collins et al. (1999), they regress stock prices on earnings and book values and then compare the coefficients for German and French firms. The results show that in the first year after IFRS adoption the coefficients do not significantly differ. In contrast, the examination of the following two years shows a significant difference in the coefficients estimated for German and French firms. The authors interpret their findings as indicative of earnings and book values being comparable in the first year after IFRS adoption, but becoming less so in the following years. They argue that institutional differences create incentives for managers to implement IFRS differently over time. Kang’s (2012) discussion of Liao et al.’s (2012) study emphasizes that the analysis could be improved by investigating the driver of comparability in the first year after IFRS adoption and by questioning why reporting incentives play a more important role in subsequent years.

3.1.2. Benefits of IFRS adoption linked to accounting comparability

Neel (2017) examines the joint effect of reporting quality and comparability on capital-market variables. The study hypothesizes and shows that firms with high reporting quality experience capital-market benefits after they adopted IFRS. Neel’s (2017) measures for capital-market outcomes (‘economic outcomes’) are Tobin’s Q (as a proxy for firm value) and the proportion of trading days with zero daily stock returns (for illiquidity), as well as analyst forecast errors and

forecast dispersion (for information asymmetry). The comparability measures used are the similarity of accounting functions metric from De Franco et al. (2011) and two of the accounting system comparability metrics employed by Barth et al. (2012). Neel (2017) finds that firms with a larger improvement in comparability experience larger increases in firm value and liquidity and larger decreases in information asymmetry relative to other adopters. In contrast, improvements in reporting quality only have a positive effect on firm value, which is moreover limited to those adopters with concurrent improvements in comparability. Neel (2017) interprets these results as being indicative of capital-market benefits after IFRS adoption being restricted to firms that exhibit an increase in comparability. He concludes that improvements in reporting quality around IFRS adoption seem to have a second-order effect, while comparability increases have a first-order effect.

Unlike many of the studies included in our survey, DeFond et al. (2011) are not principally concerned with the comparability of accounting standards. They are instead interested in real effects due to changes in foreign mutual fund ownership that could follow an increase in comparable reporting after the introduction of mandatory IFRS reporting requirements. They hypothesize that substantial increases in the uniformity¹⁰ of accounting standards attract more foreign direct investment in countries with credible implementation mechanisms in place. DeFond et al. (2011) measure changes in the degree of uniformity as the industry-specific ratio between the number of industry peers uniformly using IFRS in 2007, i.e., at one point in time after the mandatory IFRS regime had become effective, and the number of industry peers using the same (local) accounting standard in 2003, i.e., at one point in time within the pre-IFRS reporting period.

¹⁰ It may be noted that the concept of uniformity and its use in DeFond et al.'s (2011) study is remarkably different to the *de facto* comparability focus of other studies presented in our survey. While the majority of studies discussed in section 3.1.1 investigate, e.g., *whether* IFRS adoption is associated with an increase in *de facto* comparability, DeFond et al. (2011) *assume* that common accounting standards make comparative assessments easier for international investors. A similar assumption is incorporated in the research designs of Horton et al. (2013) as well as Young and Zeng (2015). As other studies show that *de jure* comparability does not necessarily lead to *de facto* comparability, this assumption could be questioned. However, since DeFond et al. (2011) conduct analyses using the similarity of accounting functions measure in their robustness checks, they effectively hedge against such criticism.

Following this measurement concept, industries with few (national) industry peers prior to IFRS adoption and many (international) industry peers after IFRS adoption exhibit a remarkable increase in uniformity.¹¹ The main empirical results show that cross-border investments did indeed increase for companies that exhibited a large boost in uniformity at the industry-level and that are subject to credible IFRS implementation at the country-level.

A study by Horton et al. (2013) deals with the change to the information environment upon the introduction of mandatory IFRS reporting regimes. Focusing on analysts' forecast accuracy (and, in sensitivity analyses, on other financial analyst variables), Horton et al. (2013) hypothesize and find that mandatory IFRS adoption is accompanied by both comparability and information quality benefits. Their measurement of comparability is based on changes to analysts' firm portfolios and predictions for different groups of portfolio changes. More specifically, analysts' portfolio changes are classified into 'Local GAAP to IFRS', 'Multiple GAAP to IFRS', and 'Local GAAP to Multiple GAAP' changes, thereby addressing the way in which the uniformity of financial reporting changed following the introduction of the mandatory IFRS regime. In the second group of analyst portfolio changes, i.e., 'Multiple GAAP to IFRS', a majority of firms initially reports under different local reporting regimes and then files IFRS reports after the IFRS mandate became effective. Horton et al. (2013) argue that comparability is likely to increase for these firms. Accordingly, for analysts experiencing such portfolio changes due to the introduction of IFRS reporting requirements, they expect and find forecast accuracy to significantly increase.

Brochet et al. (2013) test the indirect capital-market benefits that IFRS adoption brought about through the channel of enhanced accounting comparability. Capital-market benefits are measured as the reduction in abnormal returns to insider purchases. The analysis focuses on a sample of UK firms around IFRS adoption. The authors argue that UK GAAP and IFRS had

¹¹ However, as the calculated ratio of changes in the number of industry peers under the same accounting standards regime could become overly large for firms with very few peers prior to IFRS adoption, a dichotomous variable that indicates whether the increase in uniformity lies above or beneath the median change in uniformity (which is calculated at the industry-country level) is introduced and employed in their multivariate regression analyses.

negligible differences and that IFRS adoption would therefore provide a natural experiment to test the effect of enhanced comparability across a larger set of firms in the absence of confounding effects due to changes in the quality of the accounting standards. The main results show that IFRS adoption leads to significant capital market benefits. Brochet et al.'s (2013) analysis also uses three measures of comparability (the approach proposed by DeFond et al. 2011; a cross-country modification of the measurement by De Franco et al. 2011; the degree of information transfer also used by Yip and Young 2012) and documents that capital-market benefits exist for firm that experience increases in comparability after IFRS adoption.

3.1.3. IFRS adoption and comparability: Common themes and suggestions for future research

The results of most papers on the effects of IFRS adoption on comparability—in particular, Barth et al. (2012 & 2013), Yip and Young (2012), and Caban-Garcia and He (2013)—document significant increases in comparability after IFRS adoption. However, interestingly, two studies suggest that the comparability increase after IFRS adoption may be driven by a sub-sample of firms or that IFRS adoption may even be detrimental to accounting comparability. Specifically, Cascino and Gassen (2015) find that the increase in comparability is marginal and only limited to firms with distinct compliance incentives; Lang et al. (2010) even find that comparability decreases with IFRS adoption.

Future research could help to reconcile the existing evidence on the effect of IFRS adoption on comparability. It would be interesting to shed more light on the circumstances under which IFRS adoption does not result in increased accounting comparability. For example, researchers may extend the work of Cascino and Gassen (2015), who emphasize the role of compliance for comparability. Given that in the meantime the availability of data since IFRS adoption has substantially increased relative to most of the studies conducted on this topic, prospective research could also examine the pattern of accounting comparability over time after IFRS adoption and whether the increase in comparability is sustained in the medium or long term. This is related to

the early evidence reported by Liao et al. (2012), who find that comparability decreases after the first year of IFRS adoption.

A second common research question in the literature is how comparability is related to other benefits of IFRS adoption. Existing research tends to find that IFRS brought about benefits through the channel of increased comparability. DeFond et al. (2011) find that IFRS adoption leads to higher cross-country investments through increased comparability. Horton et al. (2013) find that IFRS adoption leads to higher forecast accuracy through increased comparability. Brochet et al. (2013) find evidence consistent with the notion that, after IFRS adoption, insiders' ability to profit from private information decreases through the channel of increased comparability. Neel (2017) finds that firms experiencing larger capital market benefits (in terms of liquidity, firm value, and information asymmetries) are associated with larger increases in comparability after IFRS adoption.

While the results of these studies converge, there are substantial differences in the measures of comparability used. Brochet et al. (2013) and Neel (2017) base their analyses on the measures introduced by De Franco et al. (2011) or by modifications of them. DeFond et al. (2011) and Horton et al. (2013) concentrate on the uniformity of reporting. Specifically, the observation that two firms report under the very same (and not just an arguably more similar) accounting regime, is interpreted as an improvement in comparability (see our footnote 10 for a more detailed description of this difference). Hence, the measures used by Horton et al. (2013) as well as DeFond et al. (2011) can clearly be distinguished from those of De Franco et al. (2011). Future research could extend the results on the relation between the benefits of IFRS adoption and accounting comparability. A potentially promising area of analysis is represented by unlisted firms, whose IFRS adoption consequences have been examined only by a limited number of prior studies (e.g., Gross 2016). As mentioned earlier, the measure developed by Kim et al. (2013) can be

potentially used for firms which do not have listed equity; as described in more detail in Section 2.2.2.4, this measure could be modified to obtain a firm-year (or quarter-year) specific indicator of comparability.

3.2. Determinants and consequences of comparability

Enhancing the understanding of determinants and the consequences of accounting comparability is not only highly relevant for standard setters that introduce or evaluate accounting regulation but also to financial statement users and preparers. Since accounting comparability has been consistently found to affect the quality of the information environment (starting from De Franco et al. 2011), it is likely to play an important role in affecting their choices. We group the non-IFRS-related studies on comparability into two categories: those investigating the determinants (and the time variation) of comparability and those focusing on its consequences. We note that the most influential studies on the consequences of comparability are those conducted by De Franco et al. (2011; on the quality of the information environment), Bhojraj and Lee (2002; on stock price valuation accuracy), and Kim et al. (2013; on the cost of public debt), which are already discussed in sections 2.2.2.1, 2.2.2.3, and 2.2.2.4, respectively.

All studies related to the determinants or consequences of accounting comparability that are surveyed in our paper are listed and briefly described in Table 2.

< Insert Table 2 about here >

3.2.1. Determinants of comparability

Land and Lang (2002) provide evidence of accounting standards becoming more similar across countries over time. Similar to De Franco et al. (2011), they note that the concept of (*de facto*) comparability (and any measurement of it) has to capture more similar accounting practices, while economic factors need to be held constant. However, Land and Lang (2002) operationalize the identification of comparability effects differently, employing a more ‘indirect approach’. This means that the effects of country dummy variables on (mean-adjusted) earnings-price ratios are

estimated for different countries and two different periods of time (1987–1992 and 1994–1999). The convergence of regression coefficients, i.e., an increase in the similarity of earnings multiples over time, is interpreted as an increase in accounting comparability. A plethora of additional analyses, many of which are targeted at controlling for economic trends, confirm the robustness of the results.

In line with Land and Lang (2002), Beuselinck et al. (2007) study the determinants of cross-country accounting comparability over time in a sample of EU countries prior to IFRS adoption. The main measure of accounting comparability is based on a piecewise regression (by country) of accruals on cash flows and dummy variables for positive or negative cash flows—a model that was introduced by Ball and Shivakumar (2005). The focus of the analysis is on the coefficient of the interaction term between the cash flow variable and the dummy indicating positive cash flows, which can be interpreted as a country-specific partial correlation coefficient of the relation between accruals and cash flows whenever cash flows are positive. The main result presented by Beuselinck et al. (2007) is that there seemed to be a trend, which existed prior to IFRS adoption, towards a greater cross-country comparability in the relation between accruals and cash flows. The results also show that EU firms display more similar accounts when they are in similar stages of the macroeconomic business cycle. In extensions of the basic model, Beuselinck et al. (2007) find that firm-specific and country-specific reporting incentives significantly affect the accrual-cash flow relation.

Jayaraman and Verdi (2014) examine the interaction between the convergence in reporting incentives and in accounting standards in achieving cross-country accounting comparability. Their analysis extends that of Cascino and Gassen (2015), who hypothesize and find that firms with higher compliance incentives experience greater comparability increases upon IFRS adoption. Jayaraman and Verdi (2014) use the introduction of the common Euro currency in the EU as an event that is expected to lead to higher convergence in reporting incentives due to higher economic integration. They use the mandatory IFRS adoption for listed firms within the EU in

2005 as an event that led to higher convergence in accounting standards. They measure comparability similar to De Franco et al. (2011) in combination with adaptations made by Barth et al. (2012). They document two important findings. First, they find comparability increases after the introduction of the Euro, which is in agreement with the view that greater economic integration creates incentives to report more similarly in financial statements. Second, accounting comparability only increases after the mandatory IFRS adoption in Euro countries, which suggests that reporting incentives and accounting standards are complements in achieving cross-country comparability in financial reporting. This second finding is consistent with an argument by Ball (2006), who claims that—particularly if there is substantial discretion in accounting standards—a common set of standards does not result in greater comparability if convergence in reporting incentives is absent.

Francis et al. (2014) examine whether ‘audit style’ is related to accounting comparability. Specifically, they investigate whether companies audited by the same auditor display more comparable financial statements than those audited by different auditors. For this purpose, they use two different comparability measures: First, comparability is measured as negative one multiplied by the pairwise absolute firm difference in total signed reported accruals with peers being identified as all the firms belonging to the same industry. Second, earnings comovement—as introduced by De Franco et al. (2011) as a complement to their primary measure—is used. Francis et al. (2014) regress both of the previously mentioned measures of pairwise comparability on a set of control variables and a dummy for pairs of firms audited by the same ‘Big 4’¹² audit firm. Using both measures, they find comparability to be positively associated with the dummy for firm pairs with the same Big 4 auditor. A third measurement approach used in the paper considers a model where the difference in firm-specific signed total accruals is regressed on a set of control variables

¹² Firms are classified to be audited by a Big 4 auditor, if Deloitte, Ernst & Young, KPMG, or Pricewaterhouse Coopers (PwC) audited their reports.

and on fixed effects for individual auditors. The hypothesis that auditor-fixed effects are jointly zero is rejected, suggesting that audit style plays a significant role in determining comparability.

Gross (2016) examines the *de facto* accounting comparability between private German firms that mandatorily report under local GAAP and similar firms that voluntarily report under IFRS before and after a major accounting standards reform ('German Accounting Law Modernization Act'). This accounting standards reform aligned parts of local GAAP reporting (such as provisions, deferred taxes, intangible assets, or inventory valuation) to IFRS but did not introduce full IFRS reporting. The findings, which show that comparability substantially increases after the reform, are consistent with the view that even a modest alignment of two accounting systems can foster more comparable financial reporting choices. However, since the results of Gross (2016) seem to depict a specific equilibrium for firms under various incentives, it remains questionable whether and to which extent his results are generalizable.

3.2.2. Consequences of comparability

One important expected benefit of greater accounting comparability is an improvement in the accuracy of valuations derived from comparisons of peer-pricing multiples. Young and Zeng (2015) test this conjecture in a cross-country setting around the time of IFRS adoption. They focus on the accuracy of intrinsic value estimates resulting from Bhojraj and Lee's (2002) warranted multiples method estimated for international peers. The first finding is that the accuracy of valuation did indeed increase after IFRS adoption, with the main analysis taking advantage of this increase to investigate the relation between comparability and valuation accuracy. Young and Zeng (2015) calculate the difference between the valuation error when peers are selected based on accounting data (as in the warranted multiple method) and when peers are selected by industry. They find that the difference between these two valuation errors increases after IFRS adoption, which is consistent with the view that accounting comparability is the driver of the increase in valuation accuracy. Furthermore, and again consistent with a positive association between comparability and valuation accuracy, the results show that the peers selected using the warranted

multiples method exhibit more similar economic characteristics following IFRS adoption. Finally, the extent to which reporting practices materially differed before the IFRS requirements were introduced is estimated, with the results showing that the improvement in valuation accuracy after IFRS adoption is concentrated in firms with the lowest level of pre-IFRS alignment.¹³

Fang et al. (2012) focus on the private debt market and examine the role of comparability in loan contracting. The main result is that comparability is negatively associated with the cost of debt, as measured by the loan interest spread. This finding is consistent with the notion that comparability reduces information asymmetries between borrowers and lenders in debt relationships. Further analyses by Fang et al. (2012) deal with the cross-sectional variation in the strength of the association between comparability and the cost of loan contracting. The use of more restrictive non-price terms—collateral, financial covenants, and maturity—is found to mitigate the negative association between comparability and the interest loan spread. This finding is in agreement with prior literature, which documents that price and non-price terms are traded off in debt contracting. Finally, Fang et al. (2012) examine lenders' characteristics. They find that comparability is positively associated with the number of lenders and negatively associated with the percentage of loans held by the lead lenders. This evidence reinforces the interpretation of comparability as a means of reducing information asymmetries in debt markets.

Sohn (2016) examines the association between accounting comparability and earnings management. Based on the idea that managers substitute accrual-based earnings management for real earnings management when accrual-based earnings management becomes too costly (Ewert & Wagenhofer 2005), he hypothesizes and shows that increased comparability with industry peers is associated with a decrease of accrual-based earnings management and an increase of real earnings management. For his sample that solely consists of US firms, this effect is even more pronounced in the post-SOX (the Sarbanes-Oxley Act of 2002) period than in the pre-SOX period.

¹³ Please note that this study could have also been classified as an IFRS-related study and included in section 3.1. However, since the study does not directly focus on the relation of IFRS and comparability but rather assumes that firms under IFRS are more comparable than firms under different local GAAPs, we found that the classification in section 3.2.2 is a better fit.

Sohn's (2016) bases his measurement of comparability on the primary measure used by De Franco et al. (2011).

Kim et al. (2016) study the effect of accounting comparability on perceived crash risk. They measure comparability using De Franco et al.'s (2011) main measure of comparability. Perceived crash risk is measured as the steepness of the implied volatility smirk. The results, which are based on US data, show that accounting comparability and perceived crash risk are negatively associated. This is consistent with the notion that managers' ability and incentives to withhold bad news decreases with accounting comparability; since investors anticipate this, they perceive more comparable firms as less crash prone. In line with this interpretation of the results, Kim et al. (2016) find that the association between comparability and perceived crash risk is stronger in situations when managers' ability to withhold bad news is more constrained; in particular, they focus on firms with low quality information environments, firms with weak external monitoring, and firms operating in industries with a low intensity of competition.

Chen et al. (2016) are concerned with the real effect that accounting comparability might have on acquisition decisions. Their main finding is that acquisition efficiency, measured by the profitability of the acquisition for the acquirer, is higher in cases where the comparability between the acquisition target and its industry peers is more pronounced. The idea behind this analysis is that potential acquisition targets are more easily rated in preliminary due diligence investigations, if the similarity to other firms in the same industry is higher. Interestingly, they also document that the effect of the intra-industry comparability of an acquisition target on acquisition efficiency is only present if the target firm does *not* belong to the same industry as the acquirer firm. Intuitively, accounting comparability between an acquisition target and its industry peers only becomes important if the acquirer cannot gather relevant industry experience from its own business activities and therefore needs industry peers to value the target firm against. In their analyses, Chen et al. (2016) primarily measure comparability in the same way as De Franco et al. (2011)

did; further analyses use modifications based on the asymmetric timeliness of earnings (Basu 1997) and Barth et al.'s (2012) comparability metrics.

Shane et al. (2014) investigate the association between comparability and the valuation of seasoned equity offerings (SEOs). They argue that higher accounting comparability allows underwriters to improve the assessment of firms issuing secondary equity. With greater accounting similarity to their peers, they expect the identification of firms with overvalued equity to become easier. Shane et al. (2014) indeed find that SEO firms that are more similar to their peers have lower costs of issuing new equity and experience a less severe long-run underperformance in the five years following the SEO. In further analyses, they also show that—even after controlling for real and accrual-based earnings management as well as earnings surprises—the underperformance of comparable SEO firms remains relatively low. For the measurement of comparability, Shane et al. (2014) employ a modification of De Franco et al.'s (2011) measurement.

3.2.3. Determinants and consequences of comparability: common themes and suggestions for future research

A handful of papers are concerned with the determinants of comparability and with events other than IFRS-adoption that led to a change in comparability.¹⁴ Contributions on the trend in comparability prior to IFRS adoption are provided by Land and Lang (2002) and by Beuselinck et al. (2007). Both studies find an upwards trend in cross-country comparability, although Beuselinck et al. (2007) also identify firm-specific and country-specific variables affecting comparability and its variation over time. The results of Gross (2016) indicate that also a modest alignment of two accounting systems can increase comparability. Jayaraman and Verdi (2014) document that con-

¹⁴ Of course, most studies that we discussed earlier on IFRS adoption could also be classified as studies on determinants or consequences of comparability, dependent on which role comparability played in the respective research designs. However, we isolated IFRS adoption as a common theme of many studies that we deemed worthwhile of being discussed in a separate section of our study.

vergence in incentives and accounting standards are complements in achieving cross-country accounting comparability. Francis et al. (2014) find that within-country comparability is affected by the similarity in audit styles.

We believe that it would be interesting for future research to shed more light on the determinants of accounting comparability. An important issue in measuring accounting comparability, as mentioned in section 2.2.2.1, is to disentangle economic comparability and accounting comparability; this is a problem shared by most output-based measures of accounting comparability. Researchers could, e.g., overcome this issue by examining accounting comparability around exogenous changes in economic comparability. Similar to prior research on the attributes of financial reporting quality (see, for example, Dechow et al. 2010, for a comprehensive survey), future research may also examine how governance control mechanisms and equity market incentives affect accounting comparability. Francis et al. (2014) provide a contribution on the role an economic institution can play as a determinant of accounting comparability. Relatedly, another area of future research could involve examining how the interplay of economic institutions and the characteristics of the accounting systems shapes accounting comparability.

Another group of papers examines the consequences of accounting comparability, with most of the studies finding benefits associated with comparability being more pronounced. De Franco et al. (2011) document that firms with more comparable financial statements display higher analysts' forecast accuracy, have a greater number of analysts following them, and have lower forecast dispersion among the analysts' earnings forecasts. Both Bhojraj and Lee (2002) and a closely related paper by Young and Zeng (2015) find that more comparable financial statements lead to higher stock price valuation accuracy. Furthermore, comparability has been found to be negatively associated with the cost of public (Kim et al. 2013) and private debt (Fang et al. 2012). Sohn (2016) investigates the relation between accounting comparability and earnings man-

agement, while Kim et al. (2016) examine the association between perceived crash risk and comparability. Fewer studies examine real effects of comparability. Chen et al. (2016), e.g., examine circumstances where comparability enhances acquisition efficiency.

Most papers on the effects of accounting comparability concentrate on the point of view of equity investors. Fang et al. (2012) and Kim et al. (2013) are the only contributions on the effect of comparability on debt markets; these papers consider private and public debt, respectively. Future research could investigate at greater length the effect of accounting comparability on debt markets. It would be useful, for example, to examine the role of accounting comparability in situations where the interests of the debt holders and those of the equity holders do not converge. Relatedly, it would be interesting to investigate the effects of the measure of accounting comparability developed by Kim et al. (2013) in equity markets and compare the results to those originally obtained in the debt markets. We believe it would also be interesting to explore in more detail the real effects of accounting comparability; Chen et al. (2016) provide the only contribution in this area yet. In addition, a potentially promising area for future research would be to examine situations where accounting comparability has detrimental effects on the informativeness of financial statements; this is related to the study of Lang et al. (2010), who note that comparability may have adverse effects if it implies that dissimilar events are treated similarly.

4. Conclusions

In this paper, we survey and compare output-based methods of measuring accounting comparability in empirical financial accounting research. In doing so, we provide two distinct contributions to the literature. Our first contribution is a methodological one. After introducing our survey by discussing the practical relevance of the concept of accounting comparability as well as its relevance and evolution in research in section 2, we describe and comparatively review the four studies that introduced or refined the main output-based measures of accounting comparability. Dependent on the research setting, different measures could be more or less useful, and all of the metrics come with potential advantages and disadvantages; by listing, discussing and comparing

these advantages and disadvantages, our study aims at assisting researchers to adequately choose measurement in future studies on accounting comparability.

Summing up the most important advantages of different comparability measures, the measurement approach in De Franco et al.'s (2011) study is, e.g., an appealing method to investigate firms with listed equity, which is based on generally widely available data. Yip and Young's (2012) adaptation of De Franco et al.'s (2011) measurement is useful for a refined comparability measurement that also considers that, next to the similarity facet that is more obvious, also a difference facet is inherent in comparability; furthermore, Yip and Young (2012) take a first step to transfer De Franco et al.'s (2011) measurement idea to a cross-country setting. Further improving comparability measurement in multi-country settings, Cascino and Gassen (2015) more rigorously control for confounding events in an international panel sample. Bhojraj and Lee (2002) deal with comparability from a different angle by adopting a valuation perspective on identifying comparable firms; this perspective could both be useful in studies that directly focus on comparability as well as for matching firms in other empirical studies. Finally, with their application in a bond market setting, Kim et al. (2013) provide an interesting alternative for the comparability measurement of firms in public debt markets.

Even though we consider the advantages of output-based comparability measurement to outweigh its disadvantages, we also identify some drawbacks to this measurement approach. We discuss the limitations of each of the main four output-based measures in section 2. One of the most important problems, which is common to all the output-based measures, is that they are also affected by economic comparability, i.e., the similarity with which the cash flows of the company react to economic events. Disentangling accounting and economic comparability remains a challenge from a conceptual and empirical perspective that may restrict the interpretation of respective findings. Relatedly, most output-based comparability measures are based on the assumption that economic comparability is constant across the firms within an industry-year group. However,

recent research (e.g., Srivastava 2016) shows that this assumption may not hold because of systematic differences in the production functions; moreover, the accuracy of the commonly used industry classifications is a controversial issue (e.g., Hoberg & Philips 2016).

Our second contribution is based on the content of contemporary studies on comparability in empirical financial accounting, since we provide an early survey and propose a classification for studies that make use of output-based accounting comparability measurement. We classify the existent research into three categories. The first group examines the relation between (the introduction of) IFRS and comparability (see sections 3.1.1 and 3.1.2). We classify the other studies based on their focus on the determinants or consequences of comparability. If the comparability measures are primarily used as dependent variables, we classify the respective study as being concerned with the determinants of accounting comparability (see section 3.2.1). If the comparability metrics are mainly used as explanatory variables, we interpret the respective studies as being focused on the consequences of accounting comparability (see section 3.2.2).

Within each of the content-based sub-streams of literature, we raise several suggestions for future research, particularly in sections 3.1.3 and 3.2.3. For the IFRS-related studies, we first suggest that future research should try to reconcile existing evidence to clarify when and under which circumstances IFRS adoption—or more generally the adoption of a common set of accounting standards in different countries under varying institutional environments and incentives—may be detrimental to accounting comparability. Second, researchers could also focus on the question of how different effects that may be brought about by the common IFRS adoption are linked to each other; since comparability likely mediates other effects that follow accounting standards adoption, the clarification of its role within these different effects should be of particular interest. In the other areas of research that we review, we propose to further examine the characteristics of the institutional environment that drive accounting comparability and how these determinants of comparability interact under varying incentives. When it comes to the consequences of accounting comparability, we believe that potentially interesting and scantily explored areas of

research are represented by debt market effects and the real effects that increased or decreased comparability could bring about.

By providing the methodological and content-driven contributions just described, our paper is mainly aimed at assisting empirical researchers in the field of financial accounting in comparatively assessing different measurement concepts, selecting empirical proxies for prospective research as well as identifying relevant research questions that have not yet been investigated. However, we believe that our paper is also relevant to accounting regulators. Notably, both the IASB and the FASB, e.g., are interested in accounting comparability and, accordingly, list comparability among the desirable characteristics of financial accounting information in their Conceptual Frameworks. Hence, these and other standard setters as well as enforcement bodies may be interested in how to adequately measure accounting comparability and interpret studies that aim at doing so, for example to evaluate the effect of new accounting standards or enforcement actions. Our discussion of the findings on the economic determinants of accounting comparability can also help accounting regulators to proactively design policies which ought to improve the comparability of accounting information. In addition, we review the results of many studies which focus on how capital markets are affected by accounting comparability; this part of our study is likely to be useful for accounting regulators that examine whether increasing accounting comparability achieves the overarching goal of accounting standard setters in the last decades, namely improving the decision usefulness of accounting information.

Literature

- Ali, M. J. (2005). A Synthesis of Empirical Research on International Accounting Harmonization and Compliance with Financial Reporting Standards. *Journal of Accounting Literature*, 24, 1–52.
- Baker, R. C., & Barbu, E. M. (2007a). Evolution of research on international accounting harmonization: a historical and institutional perspective. *Socio-Economic Review*, 5, 603–32.
- Baker, R. C., & Barbu, E. M. (2007b). Trends in research on international accounting harmonization. *The International Journal of Accounting*, 42, 272–304.
- Ball, R. (2006). IFRS: Pros and cons for investors. *Accounting and Business Research*, 36 (Supplement 1), 5–27.
- Ball, R., & Shivakumar, L. (2006). The role of accruals in asymmetrically timely gain and loss recognition. *Journal of Accounting Research*, 44, 207–42.
- Barth, M., Landsman, W., Lang, M., & Williams, C. (2012). Are IFRS-based and US GAAP-based accounting amounts comparable? *Journal of Accounting and Economics*, 54, 68–93.
- Barth, M., Landsman, W., Lang, M., & Williams, C. (2013). *Effects on Comparability and Capital Market Benefits of Voluntary Adoption of IFRS by US firms: Insights from Voluntary Adoption of IFRS by Non-US Firms*. Working paper. Retrieved from the Social Science Research Network (SSRN): <http://ssrn.com/abstract=2196247>
- Basu, S. (1997). The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics*, 24, 3–37.
- Beuselinck, C., Joos, P., & van der Meulen, S. (2007). *International Earnings Comparability*. Working paper. Retrieved from the SSRN: <http://ssrn.com/abstract=1014086>
- Bhojraj, S., & Lee, C. M. C. (2002). Who Is My Peer? A Valuation-Based Approach to the Selection of Comparable Firms. *Journal of Accounting Research*, 40, 407–39.
- Bhojraj, S., Lee, C. M. C., & Oler, D. K. (2003). What's My Line? A Comparison of Industry Classification Schemes for Capital Market Research. *Journal of Accounting Research*, 41, 745–74.
- Brochet, F., Jagolinzer, A. D., & Riedl, E. (2013). Mandatory IFRS Adoption and Financial Statement Comparability. *Contemporary Accounting Research*, 30, 1373–400.
- Caban-Garcia, M. T., & He, H. (2013). Comparability of Earnings in Scandinavian Countries: The Impact of Mandatory IFRS Adoption and Stock Exchange Consolidations. *Journal of International Accounting Research*, 12, 55–76.
- Cascino, S., & Gassen, J. (2015). What Drives the Comparability Effect of Mandatory IFRS Adoption? *Review of Accounting Studies*, 20, 242–82.
- Chen, C.-W., Collins, D. W., Kravet, T., & Mergenthaler, R. D. (2016). *Financial Statement Comparability and the Efficiency of Acquisition Decisions*. Working paper. Retrieved from the SSRN: <http://ssrn.com/abstract=2169082>
- Collins, D., Pincus, M., & Xie, H. (1999). Equity valuation and negative earnings: The role of book value of equity. *The Accounting Review*, 74, 29–61.
- Daske, H., Hail, L., Leuz, C., & Verdi, R. (2013). Adopting a Label: Heterogeneity in the Economic Consequences Around IAS/IFRS Adoptions. *Journal of Accounting Research*, 51, 495–547.
- Dechow P., Ge, P., & Schrand C. (2010). Understanding Earnings Quality: A Review of the Proxies, Their Determinants and Their Consequences. *Journal of Accounting and Economics*, 50, 344–401.
- De Franco, G., Kothari, S. P., & Verdi, R. S. (2011). The Benefits of Financial Statement Comparability. *Journal of Accounting Research*, 49, 895–931.

- DeFond, M., Hu, X., Hung, M., & Li, S. (2011). The impact of mandatory IFRS adoption on foreign mutual fund ownership: The role of comparability. *Journal of Accounting and Economics*, 51, 240–58.
- Ewert, R., & Wagenhofer, A. (2005). Economic Effects of Tightening Accounting Standards to Restrict Earnings Management. *The Accounting Review*, 80, 1101–24.
- Fang, X., Li, Y., Xin, B., & Zhang, W. (2012). Accounting Comparability and Bank Loan Contracting. Working paper. Retrieved from the SSRN: <http://ssrn.com/abstract=2138202>
- Francis, J. R., Pinnuck, M., & Watanabe, O. (2014). Auditor Style and Financial Statement Comparability. *The Accounting Review*, 89, 605–33.
- Gross, C. (2016). The Effect of the German Accounting Law Modernization Act on the Comparability of Private Local GAAP and IFRS Firms. *Schmalenbach Business Review*, 17, 423–60.
- Hoberg, G., & Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124, 1423–1465.
- Horton, J., Serafeim, G., & Serafeim, I. (2013). Does Mandatory IFRS Adoption Improve the Information Environment? *Contemporary Accounting Research*, 30, 388–423.
- Jayaraman, S., & Verdi, R. (2014). *Are Reporting Incentives and Accounting Standards Substitutes or Complements in Achieving Accounting Comparability?* Working paper. Retrieved from the SSRN: <http://ssrn.com/abstract=2428263>
- Kang, T. (2012). Discussion of: The Cross-Country Comparability of IFRS Earnings and Book Values: Evidence from France and Germany. *Journal of International Accounting Research*, 11, 185–90.
- Kim, S., Kraft, P., & Ryan, S. (2013). Financial statement comparability and credit risk. *Review of Accounting Studies*, 18, 783–823.
- Kim, J. B., Leye, L., Lu, L. Y., Yu, Y. (2016). Financial statement comparability and expected crash risk. *Journal of Accounting and Economics*, 61, 294–312.
- Kvaal, E., & Nobes, C. (2012). IFRS Policy Changes and the Continuation of National Patterns of IFRS Practice. *European Accounting Review*, 21, 343–71.
- Land, J., & Lang, M. H. (2002). Empirical Evidence on the Evolution of International Earnings. *The Accounting Review*, 77, 115–33.
- Lang, M. H., Maffett, M. G., & Owens, E. L. (2010). *Earnings Comovement and Accounting Comparability: The Effects of Mandatory IFRS Adoption*. Working paper. Retrieved from the SSRN: <http://ssrn.com/abstract=1676937>
- Liao, Q., Sellhorn, T., & Skaife, H. A. (2012). The Cross-Country Comparability of IFRS Earnings and Book Values: Evidence from France and Germany. *Journal of International Accounting Research*, 11, 155–84.
- Neel, M. (2017). Accounting Comparability and Economic Outcomes of Mandatory IFRS Adoption. *Contemporary Accounting Research*, 34, 658–690.
- Nobes, C. W. (Ed.). (2001). *GAAP 2001: A survey of National Accounting Rules Benchmarked against International Accounting Standards*. Retrieved from Deloitte's IAS Plus website: <http://www.iasplus.com/en/resources/binary/resource/gaap2001.pdf>
- Ohlson, J. (1995). Earnings, Book Values, and Dividends in Equity Valuation. *Contemporary Accounting Research*, 11, 661–87.
- Shane, P. B., Smith, D. B., & Zhang, S. (2014). *Financial Statement Comparability and Valuation of Seasoned Equity Offerings*. Working paper. Retrieved from the SSRN: <http://ssrn.com/abstract=2372965>
- Sohn, B. C. (2016). The Effect of Accounting Comparability on Earnings Management. *Journal of Accounting and Public Policy*, 35, 513–39.
- Srivastava, A. (2016). A Systematic Problem in the Detection of Abnormal Acts with Industry-Based Models. Working paper, Dartmouth College.

- Tay, J. S. W., & Parker, R. H. (1990). Measuring International Harmonization and Standardization. *Abacus*, 26, 71–88.
- van der Tas, L. G. (1988). Measuring Harmonisation of Financial Reporting Practice. *Accounting and Business Research*, 18, 157–69.
- Yip, R. W. Y., & Young, D. (2012). Does Mandatory IFRS Adoption Improve Information Comparability? *The Accounting Review*, 87, 1767–89.
- Young, S., & Zeng, Y. (2015). Accounting Comparability and the Accuracy of Peer-Based Valuation Models. *The Accounting Review*, 90, 2571–2601.
- Zeff, S. A. (1993). The Politics of Accounting Standards. *Academia Italiana di Economia Aziendale*, 12, 123–42.
- Zeff (2012), S. A. (2012). The Evolution of the IASC into the IASB, and the Challenges it Faces. *The Accounting Review*, 87, 807–37.

Appendix: Tables

Table 1
Overview of studies on the association between accounting comparability and IFRS

Study	Comparability measurement*	Sample period		Within- and/or between country comparability	Geographical focus	Focusing on mandatory or voluntary IFRS adoption	Content
		pre-IFRS	post-IFRS				
Barth et al. (JAE, 2012)	ASC, VRC	1992–2005	1995–2009	between	27 countries all over the world	Both (however, emphasis on mandatory adoption)	The study focuses on effects that mandatory IFRS adoption has on the comparability of IFRS and matched US GAAP firms.
Barth et al. (2013)	ASC, VRC	1996–2008		within	27 countries all over the world	voluntary	The study examines within-country comparability effects of voluntary IFRS adoption to comment on whether it could be beneficial to allow for voluntary IFRS adoption in the US.
Brochet et al. (CAR, 2013)	DKV1	2003–2004	2005–2006	within	UK	mandatory	This paper uses an UK-setting to separate the comparability effects of mandatory IFRS adoption from effects associated with potentially

							contemporaneous improvements in earnings quality.
Caban-Garcia & He (<i>JJAR</i> , 2013)	LL, DKV1	2001–2004	2005–2008	between	Denmark, Finland, Norway, and Sweden	mandatory	The authors study whether the comparability effects after IFRS adoption are more pronounced in countries that were part of the merger of Scandinavian stock exchanges.
Cascino & Gassen (<i>RAS</i> , 2015)	DKV1, CG	2001–2004	2005–2008	between	29 adopting and non-adopting countries all over the world; in-depth compliance analyses on Germany and Italy	mandatory	Controlling for time-trends by including non-adopting countries into the sample, this paper shows that an IFRS requirement does not automatically lead to higher between-country comparability if compliance incentives vary.
DeFond et al. (<i>JAE</i> , 2011)	DHHL, DKV1	2003–2004	2006–2007	neither (comparability is assumed to be induced by uniformity of reporting requirements)	14 countries in the EU	mandatory	It is tested whether a large increase in industry peers after the introduction of IFRS requirements in the EU is associated with a concurrent increase in foreign mutual fund ownership for countries with different

							implementation credibility.
Horton et al. (CAR, 2013)	HSS	2001–2004 (2001–2002 for Singapore)	2005–2007 (2003–2007 for Singapore)	neither (comparability is assumed to be induced by uniformity of reporting requirements)	45 countries all over the world (46 if Hong Kong is interpreted to be a country)	both (however, emphasis on mandatory adoption)	Horton et al. (2013) find improvements in the information environment after mandatory IFRS adoption and isolate information and comparability benefits as drivers for these results.
Lang et al. (2010)	DKV1, DKV3	2001–2004	2005–2008	between	47 countries all over the world	mandatory	Lang et al. (2010) examine changes in cross-country comparability upon mandatory IFRS adoption and the effects on the information environment. They find that DKV1 and DKV3 are associated with different effects and that an increase in DKV3 is negatively associated with the proxies for the quality of the information environment.
Liao et al. (JIAR, 2012)	CPX	2004 (only used in within-country robustness analysis)	2006–2008	both (however, focus on between-country comparability)	French and German firms	mandatory	Liao et al. (2012) find that French and German firms have comparable earnings and book

							values in the year after mandatory IFRS adoption, but this comparability effect does not prevail in subsequent years.
Neel (CAR, 2017)	DKV1, ASC	2001–2004	2005–2008	between	41 countries all over the world	mandatory	Neel (2017) hypothesizes and shows that IFRS adoption leads to capital-market benefits, in terms of valuation, liquidity and information asymmetry, for firms exposed to an increase in cross-country comparability.
Yip & Young (TAR, 2012)	DKV1, DIT, OHL	2002–2004	2005–2007	both	17 European countries	mandatory	Yip and Young (2012) analyze within- and between country comparability effects of mandatory IFRS adoption and find increased cross-country comparability. They differentiate between the similarity and the difference facet inherent in comparability by comparing firms

							from similar and dissimilar industries.
--	--	--	--	--	--	--	---

CAR stands for Contemporary Accounting Research, *JAE* for the Journal of Accounting and Economics, *JJAR* for the Journal of International Accounting Research, *RASt* for the Review of Accounting Studies, and *TAR* for The Accounting Review.

* Comparability measure abbreviations:

ASC stands for the ‘accounting system comparability’ metric introduced by Barth et al. (2012). It is based on DKV1. However, while DKV1 requires longitudinal regressions in the first step, ASC is estimated in panel regressions before and after accounting systems changed. The change in means, medians, and standard deviations of the pre-post difference is then indicating changes in the comparability of reporting practices under both accounting systems. Moreover, ASC is estimated for three different economic outcome variables used in respective regressions: stock prices, stock returns, and operating cash flows. For all these variables, ASC is calculated in six steps: First, the association between earnings and the book value of equity (earnings, change in earnings, a loss dummy, and interactions; earnings) on stock prices (stock returns; future operating cash flows) is estimated separately for firms under each accounting regime under consideration. Second, within-sample fitted stock prices (stock returns; future operating cash flows) are predicted using the regression coefficients estimated in the first-step regressions, again separately for both accounting regimes under consideration. Third, between-sample fitted stock prices (stock returns; future operating cash flows) are computed using the estimated coefficients from the first-step regressions on the book value of equity and net income (net income, the change in net income, a loss dummy, and interaction terms; net income) of firms under the other accounting regime, respectively. Fourth, again separately for firms under each accounting regime, the absolute value of the difference from fitted stock prices (stock returns; future operating cash flows) from the second and third step is calculated. Fifth, for each matched firm-year pair (i.e., one firm under each accounting regime), the differences obtained in step four are averaged. Sixth, as mentioned above, ASC is computed by using the mean, the median, and the standard deviation of the average differences obtained in the fifth step.

BL stands for Bhojraj and Lee’s (2002) warranted multiple method. Two multiples are used: the price-to-book ratio and the enterprise-value-to-sales ratio. The warranted multiples are obtained as the fitted values of yearly cross-sectional regressions of the multiples on nine explanatory variables which measure profitability, growth, and risk (the harmonic mean of the enterprise-value-to-sales-multiple for all firms in the same industry, the harmonic mean of the price-to-book-ratio for all firms in the same industry, the industry-adjusted profit margin, a variable capturing the differential effect of profit margins on the price-to-sales-ratio for loss firms, the industry-adjusted growth forecasts, the ratio of total long-term debt to the book value of common equity, the return on net operating assets, the return on equity, the ratio of research and development expenditures to sales). The comparable firms are those firms whose warranted multiple is closest to that of the target firm.

CG stands for Cascino and Gassen’s (2015) second modification of DKV1 that is on the one hand also (like their initial cross-country modification of DKV1) suitable for cross-country comparisons and on the other hand uses cash flows instead of returns to guarantee that differing levels of market efficiency are not biasing the results. A similar measure to CG is the cash flow based ASC.

CPX stands for a comparability assessment based on Collins et al. (1999) used by Liao et al. (2012). Here, stock prices are regressed on earnings per share and book values of common equity per share (together and in separate models). Comparing the coefficients of these regressions that are estimated yearly for two different groups of firms allows commenting on comparability between these groups of firms. In the study by Liao et al. (2012), this procedure is used to evaluate the comparability between French and German firms around the introduction of IFRS reporting requirements.

DHHL is a measure by DeFond et al. (2011), who use the uniformity in accounting standards to identify firms with potentially larger comparability benefits. They observe the number of industry peers before and after adoption of a uniform accounting standards regime (in their study: IFRS) and use the number of industry peers after adoption over the number of industry peers before adoption to depict the change in uniformity due to the introduction of the new accounting standards regime. In further analyses, DeFond et al. (2011) use a dichotomous variable indicating large increases in uniformity if the calculated ratio is greater than its industry-country-median.

DIT stands for the degree of information transfer employed by Yip and Young (2012). This measure is no comparability metric as such. The degree of information transfer is based on abnormal returns of non-announcing firms at earnings announcements of announcing firms. Thus, for this means of measuring comparability, no single score exists that indicates high or low comparability. However, the whole procedure allows commenting on the information transfer with respect to the news in earnings announcements.

DKV1 stands for the ‘similarity of accounting functions’ measure by De Franco et al. (2011). This measure is based on a regression of earnings (as a proxy for the accounting outcome) on returns (as a proxy for economic events) to estimate the ‘[...] mapping from economic events to financial statements’ (De Franco et al. 2011, p. 896). In the first step, this regression is estimated over the last 16 quarters for each firm individually. The regression intercept and the regression coefficient on the return variable determine the accounting function of each firm. In the second step, these accounting functions for each firm are used together with its return variable to predict earnings. Thirdly, the 16 quarterly returns of firm i are used together with the accounting functions of each other firm j in the same two-digit SIC-industry to predict earnings again. With this procedure, the economic events are explicitly held constant, while the accounting functions differ across firms. Fourth, the similarity of accounting functions between firm i and each firm j is determined by calculating the absolute difference between the two differently predicted earnings, dividing it by 16, and multiplying it by minus one. The multiplication is only done to produce greater values if comparability is high. Fifth, after calculating this similarity of accounting functions for all combinations of firms ij in each two-digit SIC-industry, firm-level measures are computed by ranking all $J-1$ similarity values to firm i and by calculating the mean of the four peer firms j exhibiting the highest comparability with firm i (*CompAcct4*) as well as the median of all $J-1$ comparability values (*CompAcctInd*).

DKV1 has arguably been the most influential comparability measure in contemporary empirical accounting research. However, for many of the questions arising in international financial accounting studies, the industry-based approach within one country is not directly suitable, since the comparability of firms situated in different countries needs to be adequately addressed. Many of the studies that deal with such questions therefore adapt DKV1 for a cross-country context. Yip and Young (2012) and Cascino and Gassen (2015), e.g., both use a cross-country modification of DKV1. Apart from heterogeneous lengths of the return calculation periods and different data structures, these two modifications of DKV1 are practically identical.

DKV2 stands for the ‘prices lead earnings’ measure by De Franco et al. (2011). DKV2 is a slight modification of DKV1. Instead of estimating the effect of stock returns on earnings in the first step, the effect of stock returns and lagged stock returns on earnings is estimated.

DKV3 stands for the ‘correlated financial statements’ measure by De Franco et al. (2011). 16 quarters of data are used to estimate a pair-wise regression of one firm’s earnings on another firm’s earnings. This effect is estimated for all firm-pair combinations within separate sample industries. DKV3 is finally defined for each firm i to be the average R^2 of the four firms js with the highest R^2 s in these regressions. In analyses containing this measure, De Franco et al. (2011) control for the correlation between operating cash flows and the correlation between stock returns.

HSS describes the comparability measurement by Horton et al. (2013). They classify analysts’ portfolio changes into different groups, addressing if uniformity of financial reporting changed after the introduction of a mandatory reporting (in their study: IFRS) regime. They expect and find comparability benefits (induced by increased consensus amongst analysts) for firms changing from a rather heterogeneous reporting environment, with firms using different accounting standards, to a rather homogenous environment, where firms report under a uniform standard.

KKR stands for the two comparability measures proposed by Kim et al. (2013). The measures are defined as minus the variability of Moody’s adjustments (within the Financial Metrics database) affecting the interest-coverage ratio or of the adjustments for non-recurring income items within an industry-peer group. To define an industry the Moody’s industry classification is used; because Moody’s adjustments are provided on a quarterly basis, the measures are calculated using quarter-industry peers. More specifically, the variability of the adjustments is measured as the difference between the upper and the lower quartile of the quarter-peer distribution of the adjustments to the interest coverage or the adjustments for the non-recurring items adjustments. The resulting measures of comparability are industry-quarter specific.

LL stands for Land and Lang’s (2002) methodology; they interpret to observe between-country comparability increases if earnings and price multiples of firms in different countries converge.

OHL stands for the ‘similarity of ICE and ICBV’ measure based on Ohlson (1995), used by Yip and Young (2012). In this model, firms’ market values are regressed on net income, the book value of equity, an industry or country indicator and the interaction of this indicator with net income and the equity book value. If the coefficient on the interaction term with net income (equity book value) is insignificant/significant, an ICE (ICBV) comparability core of 1/0 is assigned.

VRC stands for the ‘value relevance comparability’ introduced by Barth et al. (2012). Similar to ASC, it is estimated for stock prices (P), stock returns (RET), and operating cash flows (CF). It is the adjusted R² difference from the following regressions and their nested version that only include industry and country controls:

$$P_{it} = \alpha + \beta_1 \cdot BVE_{it} + \beta_2 \cdot NI_{it} + \sum_{j=1}^{J-1} (\delta_j \cdot I_{ji}) + \sum_{k=1}^{K-1} (\delta_k \cdot C_{ki}) + \varepsilon_{it}$$

$$RET_{it} = \alpha + \beta_1 \cdot \frac{NI_{it}}{P_{it-1}} + \beta_2 \cdot \frac{\Delta NI_{it}}{P_{it-1}} + \beta_3 \cdot LOSS + \beta_4 \cdot LOSS \cdot \frac{NI_{it}}{P_{it-1}} + \beta_5 \cdot LOSS \cdot \frac{\Delta NI_{it}}{P_{it-1}} + \sum_{j=1}^{J-1} (\delta_j \cdot I_{ji}) + \sum_{k=1}^{K-1} (\delta_k \cdot C_{ki}) + \varepsilon_{it}$$

$$CF_{it+1} = \alpha + \beta \cdot NI_{it} + \sum_{j=1}^{J-1} (\delta_j \cdot I_{ji}) + \sum_{k=1}^{K-1} (\delta_k \cdot C_{ki}) + \varepsilon_{it+1}$$

Here, BVE stands for the book-value of equity, NI abbreviates net income, J stands for the total number of industries I in the sample, K stands for the total number of countries C in the sample, and LOSS is a dichotomous variable that indicates negative net income values.

Table 2
Overview of studies on the determinants and consequences of accounting comparability

Study	Comparability measurement*	Comp. measure(s) as dep./indep. Variables	Other dep./indep. Variables	Sample period	Within- and/or between country focus	Geographical focus	Content
Beuselinck et al. (2007)	Country-specific measure based on a piecewise regression of accruals on cash flows	Dependent	indep.: - time; - firm-specific variables (size; leverage; labor intensity); - country-specific variables (development of the stock market; importance of the banking industry; and union membership in a country)	1990–2005	between	15 EU countries	Beuselinck et al. (2007) investigate the time variation in cross-country accounting comparability before the IFRS reporting requirements became effective.
Bhojraj & Lee (JAR, 2002)	BL	Independent	dep.: - valuation accuracy (predictive ability of the warranted multiples method)	1982–1998	within	US	Bhojraj and Lee (2002) propose a method for selecting comparable firms based on valuation theory.
Chen et al. (2016)	DKV1	Independent	dep.: - acquisition efficiency	1983–2009	within	US	Chen et al. (2016) statistically examine the effect of accounting comparability on acquisition efficiency.
De Franco et al. (JAR, 2011)	DKV1; DKV2; DKV3	independent	dep.:	1993–2007	within	US	De Franco et al. (2011) comment

			<ul style="list-style-type: none"> - analysts' choice of comparable firms; - analysts' coverage of firms pairs as similar firms; - future analysts' coverage; - analysts' forecast accuracy; - analysts' forecast dispersion 				on effects of accounting comparability on the quality of the information environment (approximated by several financial analyst variables).
Fang et al. (2012)	DKV1	Independent	dep.: <ul style="list-style-type: none"> - credit spread on private loans 	1981–2008	within	US	Fang et al. (2012) examine the effect of accounting comparability on the cost of private debt.
Francis et al. (TAR, 2014)	Pairwise difference in total signed accruals; DKV3	Dependent	indep.: <ul style="list-style-type: none"> - dummy variable for pairs of firms audited by the same auditor 	1987–2011	within	US	Francis et al. (2014) find an effect of audit style on accounting comparability.
Gross (sbr, 2016)	ASC, VRC	Dependent	indep.: <ul style="list-style-type: none"> - Dummy variable for a German accounting standards reform that partly aligns local GAAP with IFRS rules for private firms 	2003–2011 (pre-reform: 2003–2009; post: 2010–2011)	within	Germany	Gross (2016) finds that the <i>de facto</i> comparability between private firms reporting under local GAAP and under IFRS increased after an accounting standards reform became effective that partly aligned local GAAP and IFRS.

Jayaraman & Verdi (2014)	DKV1	Dependent	indep.: - economic integration events (Euro adoption, IFRS adoption)	1994–2007	between	15 EU countries	Jayaraman and Verdi (2014) document an effect of economic integration (approximated by the Euro introduction in the EU) on accounting comparability.
Kim et al. (RAST, 2013)	KKR	Independent	dep.: - traded bonds' bid-ask spread; - credit spread on traded bonds	2005–2010	within	US	Kim et al. (2013) find an effect of accounting comparability on the cost of public debt.
Kim et al. (JAE, 2013)	DKV1	Independent	dep.: - perceived crash risk	1996–2013	within	US	Kim et al. (2013) find an effect of accounting comparability on perceived crash risk.
Land & Lang (TAR, 2002)	Similarity in accounting multiples	Dependent	indep.: - time	1987–1999	between	Australia, Canada, France, Germany, Japan, UK, US	Land and Lang (2002) examine the time variation in cross-country accounting comparability before the IFRS adoption.
Shane et al. (2014)	DKV1	independent	dep.: - short- and longer-run return measures	1984–2010	within	US	Shane et al. (2014) investigate the association between accounting comparability and the assessment of firms at seasoned equity offerings

							and present results on underpricing and long-run underperformance.
Sohn (<i>JAPP</i> , 2016)	DKV1	Independent	dep.: - accrual-based earnings management; - real earnings management	1980–2009	within	US	Sohn (2016) examines the effect of accounting comparability on earnings management.
Young & Zeng (<i>TAR</i> , 2015)	IFRS adoption in quasi-natural experiment is assumed to lead to greater comparability, BL	Independent	dep.: - valuation errors using the warranted multiples method	1997–2008	between	15 EU countries	Young and Zeng (2015) find an effect of accounting comparability on valuation accuracy.

JAЕ stands for the Journal of Accounting and Economics, *JAPP* stands for the Journal of Accounting and Public Policy, *JAR* stands for the Journal of Accounting Research, *RAS* for the Review of Accounting Studies, *sbr* stands for Schmalenbach Business Review and *TAR* for The Accounting Review.

* The comparability measure abbreviations are detailed in the notes to Table 1.